AD 740 658

FINAL REPORT

PART I

ANALYSIS OF ADVANCED DATA TRANSMISSION TECHNIQUES

SEPTEMBER 1962

TO FEDERAL AVIATION AGENCY

CONTRACT NO. FAA/BRD-80

1947

GENERAL DYNAMICS | ELECTRONICS-ROCHESTER

205

PAGES ______

ARE

MISSING

IN

ORIGINAL

DOCUMENT

FINAL REPORT

PART I

ANALYSIS OF ADVANCED DATA TRANSMISSION TECHNIQUES

GENERAL DYNAMICS | ELECTRONICS-ROCHESTER

Rochester, New York

September 1962

to
FEDERAL AVIATION AGENCY

Contract No. FAA/BRD-80

"This report has been prepared by General Dynamics/Electronics-Rochester for the Systems Research and Development Service, Federal Aviation Agency, under Contract No. FAA/BRD-80. The contents of this report reflect the views of the contractor, who is responsible for the facts and the accuracy of the data presented herein, and do not necessarily reflect the official views or policy of the FAA."

PREPARED BY: G. C. Porter

G. C. Porter
Project Engineer

L. M. Luke

L. M. Luke
Section Head

APPROVED BY: T. G. Hame

T. G. Hame
Manager
Communications and Electronics
Advanced Development Engineering

TABLE OF CONTENTS - PART I

TABLE OF CONTENTS - PART I (Cont'd.)

TABLE OF CONTENTS - PART I (Cont'd.)

TABLE OF CONTENTS - PART I (Cont'd.)

LIST OF FIGURES - PART I

LIST OF FIGURES - PART I (Cont'd.)

Section

LIST OF FIGURES - PART I (Cont'd.)

LIST OF FIGURES - PART I (Cont'd.)

Section

INTRODUCTION

A description of work performed to December 30, 1960, in completing the requirements under Contract FAA/BRD-80 "Analysis of Advanced Data Transmission Techniques" is presented in the following final report. For a more detailed discussion of work completed before March 1960, refer to an earlier report: "Preliminary Report for Task I and Interim Report for Task II."

During the performance of the contract, certain of the original requirements were modified in the light of results obtained in the early stages, and certain details of the original three tasks were revised and clarified to reflect shifts in emphasis. The following paragraphs present a brief review of the major items under each task, indicating the relevant sections in the report where applicable.

The report is collated with the relevant figures and charts at the back of each section. Text material in each section is page numbered independently. In addition, the material in Sections 3 (Test Facility) and 4 (Converter Design) has been separated into two parts. Portions of these sections dealing with circuit design, layout, and operating details have been bound as Part II. Sections 3 and 4 of Part I provide brief descriptions of the test facility and converters together with a short discussion on operating and design principles.

<u>TASK I</u> As originally defined, Task I was virtually completed by March 1960, except for delivery of the Potter Instrument Company tape

printer. Delivery of this printer was delayed because of unforeseen technical difficulties, among which were belt breakage in the type wheel drive coupling and non-uniform inking and spacing of the printed characters. As of December 1960, the Potter Instrument Company had delivered one model of their airborne printer to General Dynamics/Electronics - Rochester after demonstrating satisfactory performance at printing rates up to 30 characters per second in their laboratory. This unit was subsequently delivered to NAFEC.

A major item under Task I was the provision of experimental flight test equipment including two airborne sets and one ground set. Delivery of this equipment was completed in March of 1960. Section 1 of the report describes briefly the characteristics of the flight test equipment; a more detailed discussion is available in the preliminary report.

One model 118 motorless cockpit display unit, made by the Teleprinter Corporation, was delivered with the first airborne test set in January 1960. Task I was subsequently expanded to include the procurement of one additional display unit, which was delivered to NAFEC in November 1960.

A further expansion of Task I was the development of a set of binary signalling components for phase reversal keying at 60 bauds. This equipment, discussed in detail in Section 2 of the report, consisted of two signalling and time base packages designed for retro-fit in the FSK data test sets produced earlier and was delivered to NAFEC in December 1960.

for the development of a suitable test facility and construction of a workable model in breadboard form. The design characteristics and operation of this facility are discussed in Section 3.

The results of laboratory studies on FSK signalling are presented in Section 5. Among the results reported here are measurements on the effect of pulse length; the efficiency of a method of error correction based on parity checking and single null zone detection; a qualitative indication of the relationship between fortuitous distortion and SNR in optimized converters; design parameters of optimum converters; and measurements on the influence of synchronization error on bit error rate.

TASK III Section 6 presents a discussion of certain significant factors in the design of ground-air-ground communication systems. Among the questions considered are sequencing methods, radio channel requirements, radio equipment requirements, extended message handling, pilot initiated messages, modulation, and coding. Certain pros and cons related to these factors are presented, but no attempt is made to define a system.

ACKNOWLEDGEMENTS In addition to unwitting contributions to this report on the part of those authors listed in the bibliography, the efforts of Messrs. K. H. Pomeroy, B. P. Updike, D. L. Cloonan, C. H. Koerner, and G. T. Parker of General Dynamics/Electronics - Rochester are gratefully acknowledged. Appreciation is expressed to Dorothy Olson for typing the final copy.

SECTION 1.0 DATA TEST SETS

1.1 Introduction

The following section presents a brief summary of function and operational characteristics of the Data Test Sets. An earlier report [28] presented a detailed description of the logical arrangement and circuit configuration of these sets.

1.2 Functional Description

1.2.1 General

One ground set and two airborne sets have been delivered. The two airborne sets are similar except for the inclusion of an error correction system based on parity check and null zone reception in the second set. The circuit arrangement of the second airborne set results in the rejection of all bit decisions in which a null indication occurs. In these cases, an arbitrary "mark" decision is made, and the parity check completed on the resulting code pattern. Should the parity check fail, an attempt is made to correct the error if and only if one null decision is involved; otherwise, the information is rejected.

1.2.2 The Ground Set

The ground set is designed for operation with standard teletype peripheral equipment operating at 100 words per minute (10 characters per second). Five hole punched tape is used for both input and output. A

Teletype Model 28RT set and a Model 28 Multiple Contact Reader are used as input machines, but two Model 28 readers could be used equally well for test purposes (Fig. 1.1). The ground set includes automatic control circuits for the tape reader clutches. Exact synchronization between data set and input machines is achieved through the use of the 60 cycle power frequency as a common synchronizing source (Fig. 1.2).

The readers feed five bit parallel codes into the data set, which performs a parallel to serial conversion and adds a sixth parity check bit. The resulting synchronously timed DC pulses modulate a tone shift keyer, which in turn may modulate a radio transmitter or feed a line network directly.

The receiving portions of the ground set include a demodulator for converting tone shift signals into DC pulses (Fig. 1.3). Following the converter, parity and null checks are performed, and the signal feeds via a serial to parallel converter to the output tape punch in parallel form. Failure of either the parity or null check, or both, will cause the ground set receiver to substitute and punch out distinctive error characters in place of those characters in which failures occur.

since it is assumed that all incoming signals will be closely synchronized with the ground data set transmissions, except for a fixed delay, the receiver timing pulses are derived from the 60 cycle power line but are all shifted in time from the transmitter timing pulses by a pre-determined interval which may be adjusted manually.

As will be seen in a later paragraph, the airborne set responses follow identical formats in which the 7th, 8th, and 9th characters are to be interpreted as numbers. However, the response contains no figure-letter shift information; instead, this is added into the output tape automatically by the ground data set. As a consequence, the output punch must complete 12 punch cycles for each 10 characters received. This requirement is easily met by the High Speed (60 characters per second) Teletype Tape punch which is used as the ground receiver output machine.

1.2.3 The Airborne Set

The airborne set (Fig.1.4) accepts tone shift signals and converts them to DC pulses by means of a phase sensitive discriminator circuit. Mark and space energies integrate separately on two capacitors. A synchronously timed pulse discharges both capacitors at the end of each bit, and a decision circuit produces

mark or space outputs according to the relative integrated charge on the two capacitors.

After parity and null checks are performed, the output or the decision circuit is converted to parallel form for presentation to call decoder and display control circuits. The call decoder is designed to respond to either of two five character combinations, one of which is unique to each airborne set (the private call) and one of which is common to each airborne set (the general call). Reception of either call permits the remainder of the received message to operate the airborne displays.

Two types of display are provided. Fixed format messages control illuminated alphanumeric and lamp displays by means of decoders and stores consisting of magnetically latched relays in the control box (Fig. 1.5). Extended messages are displayed by a miniature tape printer using FAA weather alphabet (Fig. 1.6).

Synchronization of the airborne receiver is accomplished in two steps. The first involves phase locking a time base oscillator to the received signal, and the second involves the phasing of a count-down chain used to obtain character and block sync pulses from the time base oscillator. This second step is accomplished by

means of a special block sync pattern, which is transmitted about 1% of the time. Synchronization during transmission is dependent on the decay time of the AFC correction loop, since no phase correction pulses are available during that time.

Transmission will occur after reception of a private call, except in two special cases. One of these occurs when a ground-to-air voice message follows the call, and the other occurs when a transmit button in the airborne set has been depressed. In this case, transmission will occur following the reception of an idling signal. A pair of contacts in the airborne set is provided to control T/R switching in the RF equipment. Depression of the transmit push button illuminates a red indicator which remains on until a transmission occurs.

Data for the air-to-ground transmission is encoded by means of five manually operated coding switches on the control box. Each switch provides for the selection of one out of ten codes.

A voice output jack on the control box connects to the radio receiver output via a latching relay and a band stop filter designed to attenuate data signal frequencies. The relay contacts in this circuit can be closed by a

data signal originated by the ground unit. A "voice" push button is provided on the control box for releasing the voice relay contacts.

1.3 Characteristics

1.3.1 Speed

Operating speed of the data sets is 100 wpm (10 characters per second) fixed by the teletype input equipment. The electronic equipment is capable of higher speeds, in the order of 100 times faster; however, the decoding and storage relays of the airborne control box require about 3 milliseconds setting time, so that speeds in excess of about 300 bits (50 characters) per second would require additional buffer storage for one character. The airborne printer produced by the Teleprinter Corporation is capable of 10 characters per second, while the Potter Instrument machine will handle up to 25 characters per second.

1.3.2 Coding

Teletype coding less start and stop pulses is used. A sixth bit is added to each code to provide odd parity check. All code pulses are of equal length so that at 100 wpm (10 characters per second) the bit rate is 60 per second and the pulse length 1/60 of a second.

1.3.3 Message Format

Since the airborne printers require parallel six bit code for their operation, it is necessary to employ usual figure/letter shift characters during the extended message part of the ground-to-air transmission. In fixed format messages, classification of the received codes into figure or letter groups is based on knowledge of the message structure, and figure/letter shift characters must be omitted.

All transmissions are based on a message block of ten characters of 60 bits (Fig. 1.7). Four block types are used:

- Fixed format message,
- Synchronization,
- Idling, and
- Extended message.

In the fixed format block, the first five characters are the address and the remaining five the information. The synchronization block is used to phase counting circuits in the airborne time base; in a set which is already synchronized correctly, it is used to operate a confidence check indicator. Use of the idling block in connection with the transmit push button of the airborne set has already been explained; this block is automatically transmitted from the ground unit

when neither message nor synchronization blocks are available.

Extended messages regardless of length must consist of an integral multiple of ten characters; hence, the extended message block is in reality a very informal structure, except that the last block in any extended message must be padded out by means of blanks to exactly ten characters.

The airborne response is not listed separately because it follows the fixed format message block exactly.

1.3.4 Timing

Block transmissions from the ground unit are continuous, with no gaps between. Airborne responses normally follow the end of the ground-to-air private call with two major delays: a fixed two-character delay and an additional delay of about one bit, due to the detection and decision process in the data receiver.

In the extended message case, the response is delayed until the end of the extened message, and in the case of a private call followed by a voice message, no response at all is allowed.

Responses to an idling block follow after the normal delays. Synchronizing and general call blocks evoke no response from the airborne set. It is not possible

to interrogate the sa[illegible]orne set in two consecutive blocks. In the first [illegible] following the first interrogation, the airborne set will normally be transmitting and the receiver will be disabled. Due to normal system delays, only a part of the second block following initial interrogation can be received. It is accordingly useless to re-interrogate the same airborne set more frequently than once every third block. These limitations do not apply when the ground-to-air message blocks belong to that group which does not evoke a response in the airborne set.

In normal operation involving repetitive roll call of more than four airborne sets, ground-air interrogations and air-ground responses may occur continuously.

1.3.5 Physical

The data test sets are designed and constructed for convenience in experimentation rather than efficiency in packaging. Circuits are hand wired on plug-in cards, which are mounted in modular trays. The circuit design employs resistor-transistor logic using relatively inexpensive transistors; this factor should be taken into account in making comparisons with devices using diode logic, since such circuits will frequently require fewer transistors but many more diodes. It should also be noted that the design approach placed more emphasis on

experimental convenience than on circuit simplification.

The sets operate on about 25 watts of primary input power, exclusive of the airborne printers, and are designed for continuous operation with DC supply voltages ranging from 24 to 32 volts. The -28 volt terminal of the airborne sets is connected to chassis.

The Teleprinter cockpit printer requires current pulses of up to 3-1/2 amperes peak at 28 volts. Average requirements depend on code configurations and duty cycle and are in the order of 1-1/2 amperes at 28 volts.

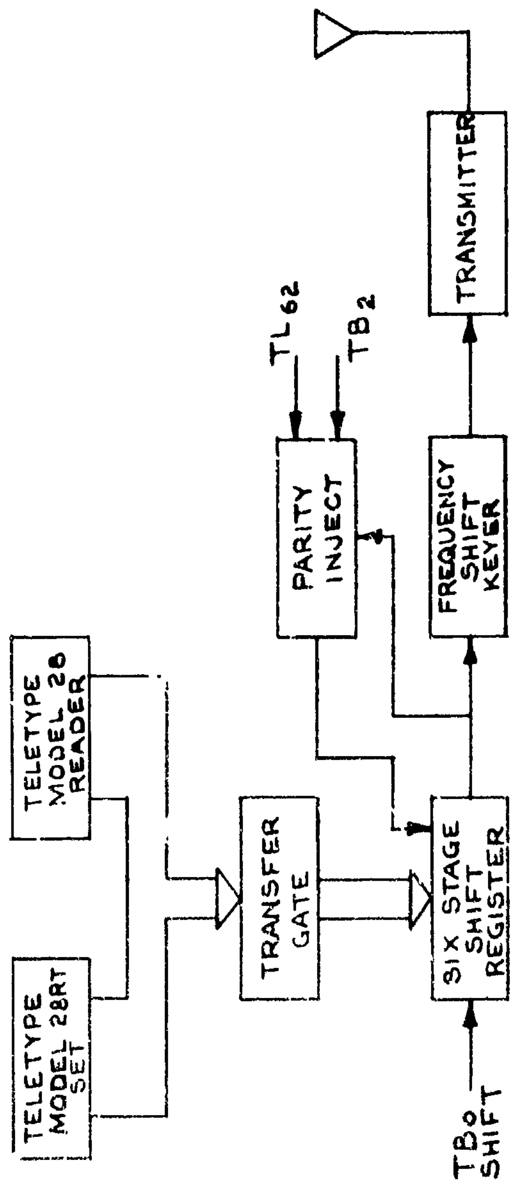
1.3.6 Signalling

Tone converter and demodulator circuits are designed for operation from 600 ohm sources at a power level of 0 dbm. Tone keyer amplifiers have a continuous adjustment to a maximum output of +15 dbm into 600 ohms. Consequently, the input signal may be taken directly from the audio distribution system of most aircraft, and the output signal may be fed into the microphone jack of most transmitters.

Tone signalling is accomplished at a center frequency of 3315 cps with shifts of $\pm$ 42.5 cps to minimize interference with voice signals. Simultaneous voice and data modulation on the ground-air signal is possible.

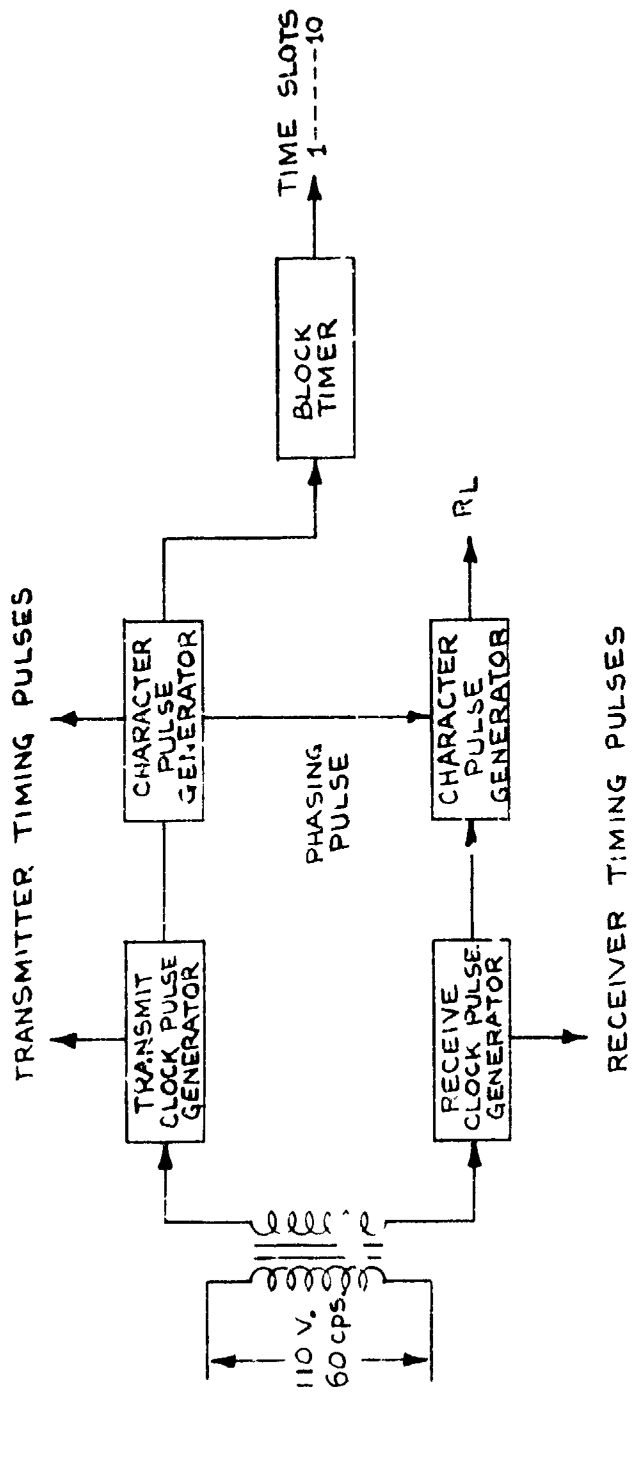
The equipment is intended for operation on dual radio

channels, where one is allocated to the ground-air transmission and the other to all air-ground transmissions. At 100 words per minute, up to sixty interrogation and response cycles per minute are possible, as well as voice operation on the ground-air path.



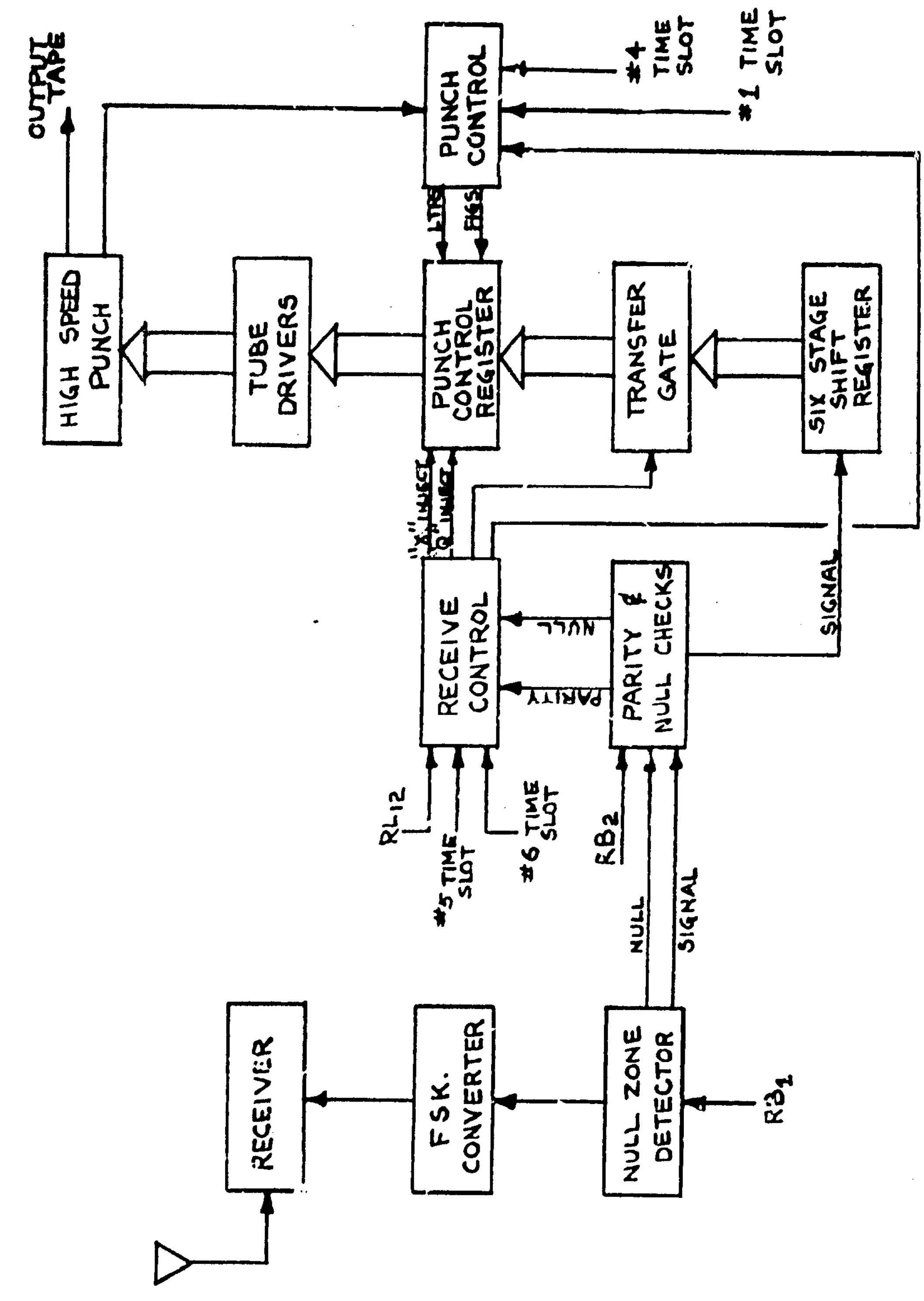
GROUND TEST SET
(TRANSMITTER)

FIG. 1.1



GROUND TEST SET
(TIMING PULSE GENERATOR)

FIG. 1.2



GROUND TEST SET (RECEIVER)

FIG. 1.3

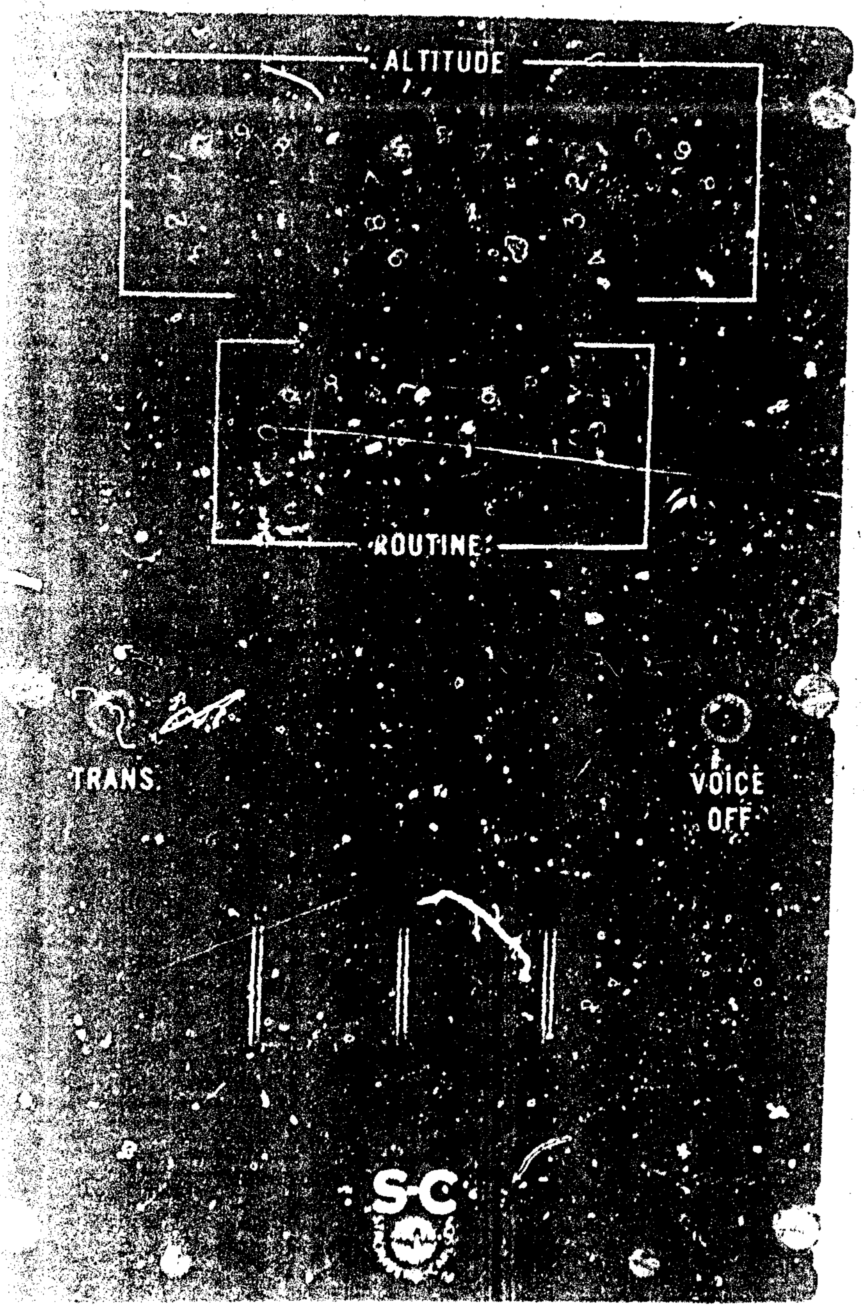
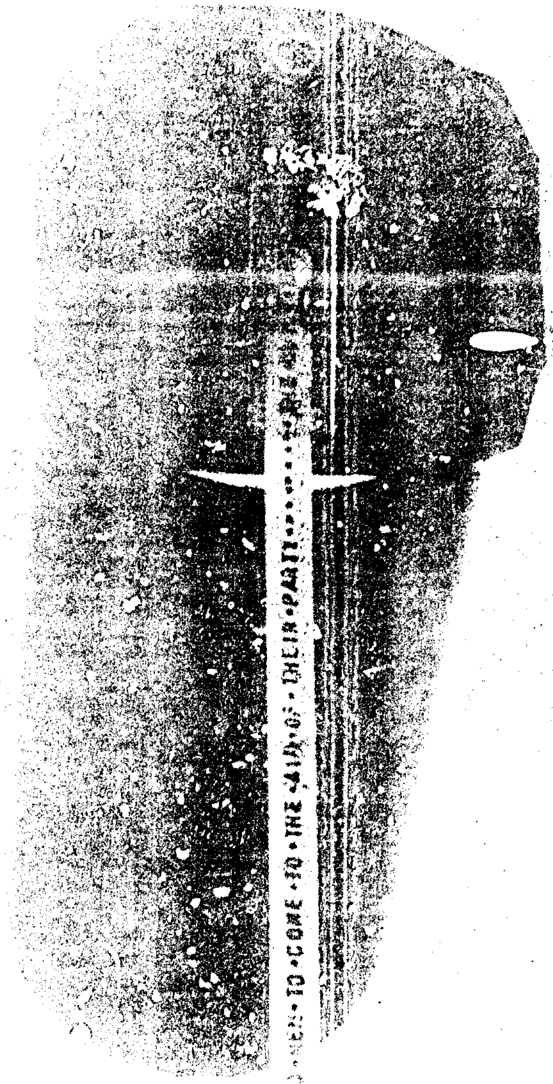


Fig. 1.[illegible]


TO • COME • TO • THE • AID • OF • THEIR • PARTY

BLOCK CHARACTER NUMBERS

1	2	3	4	5	6	7	8	9	10

FIXED FORMAT BLOCK

5 CHARACTER ALPHA-NUMERIC CALL (PRIVATE CALL) OR: R R R R W (GENERAL CALL)	B V A X	0-9	0-9	0-9	R Y G S T

SYNCHRONIZING BLOCK

R R R R W (GENERAL CALL)	R	R	R	R	BL

IDLING SIGNAL

R	Y	R	Y	R	Y	R	Y	R	Y

NOTES: 1) IN CHARACTER 6, THE SYMBOLS SHOWN APPEAR IN THE CONTROL BOX DISPLAY.

2) CHARACTERS 7 TO 9 INCLUSIVE PERMIT THE TRANSMISSION AND DISPLAY ON THE CONTROL BOX OF ANY NUMBER FROM 000 TO 999

3) IN CHARACTER 10, R, Y, AND G, OPERATE COLORED DISPLAY LAMPS IN THE CONTROL BOX, S OPERATES THE VOICE RELAY, AND T SWITCHES IN THE AIRBORNE PRINTER FOR EXTENDED MESSAGE RECORDING. NOTE THAT R, Y, AND G CODES IN THIS POSITION WILL BE IGNORED ON GENERAL CALL.

MESSAGE BLOCK FORMATS FIG. 1.7

SECTION 2.0 PHASE SHIFT KEYING

2.1 Introduction

Theoretical analyses have shown that $\pi/2$ phase modulation is superior to frequency modulation for all signal-to-noise ratios[3,20]. In this section, the results of some laboratory tests undertaken to check the practicability of exploiting this result are described. The effects of keying rate on a fixed-frequency phase-modulated sub-carrier are discussed briefly.

Design details and operating principles of a PSK signalling set suitable for retro-fit in the data sets described in Section 1 are presented.

2.2 Laboratory Tests

2.2.1 The PSK Breadboard

The PSK experimental breadboard circuit is shown in Fig. 2.1. An oscillator adjusted to operate at the center frequency of a quenched integrating filter drives a count-down circuit to generate four sets of square waves which serve as the keying signal for a phase keying circuit. Since the oscillator is common to the phase keyer and count-down circuits, coherent phase keying can be obtained with 64, 32, 16, and 8 sub-carrier cycles per pulse. A linear adder combines the keyer output with random noise and feeds the resultant into a synchronously quenched high-Q

filter. A diode phase detector compares the filtered signal with a noise-free reference frequency obtained from the oscillator, producing a DC output proportional to the phase difference and amplitudes of its inputs. An error counting circuit checks this demodulated output against the transmitted signal at the end of each bit, registering an error whenever disagreement is noted. Signal and noise powers are measured at the output of the quenched filter by a power amplifier driving a Polyranger Thermal Ammeter.

2.2.2 Calibration Check

The performance of the PSK breadboard was tested by applying unkeyed signal plus wideband noise to the quenched filter and changing the quenching rate. Ideally, the SNR at the filter output would be expected to halve when the quench rate doubles, provided that the following conditions are met:

- The noise power density is constant over the significant frequency range of the filter response.
- The ratio of signal-to-noise power applied to the quenched filter is constant.
- The measurement technique is accurate.
- No limiting or saturation occurs in the filter or measuring circuits.

The results obtained are tabulated below:

TABLE 2.1

Quench Rate in Pulses per Second	SNR in DB				
51	+6	+3	0	-3	-6
102	+2.9	0	-3.0	-5.5	-8.0

We may conclude from Table 2.1 that the equipment performs reasonably well at signal-to-noise ratios between -3 and +6 db, since the SNR ratio is degraded by approximately 3 db over this range when the quenching rate doubles. Reasons for the increasing discrepancy noted at low SNR were not investigated.

2.2.3 Effect of Pulse Length

Fig. 2.2 shows measured error rate vs. SNR at four quenching rates. The results for rates from 51 to 102 pulses per second fall within one db of each other, and when the equipment limitations are considered a difference of this magnitude is probably not significant. However, the curve for 408 pulses per second, corresponding to 8 sub-carrier cycles per bit, shows a departure from normal behavior of about 3 db. This degradation is apparently due to the width of the quenching pulse, which effectively shorts out the filter for about 1.2 milliseconds at the start of each bit. Since the signal pulse length at 8 cycles per bit is about 2.45 milliseconds, the intergrating time is only 1.25 milliseconds, and the effective bandwidth about 2.45/1.25

or about twice optimum. The 3 db performance degradation observed at 8 cycles per bit can accordingly be attributed to the width of quenching pulse. Although some improvement in performance can be expected if more care is observed in the design of the quenching circuit, it is probably fair to say that the quenching interval could not easily be made less than one sub-carrier cycle in duration, so that the minimum degradation would be about 0.6 db at 8 cycles per bit, or 1.25 db at 4 cycles per bit.

2.2.4 PSK vs. FSK

In his paper, Montgomery[3] derives the following expression for probability of bit error in PSK signalling:

$$p(e) = 1/2 \; (1 - I \sqrt{R}),$$

where $I \sqrt{R}$ is the probability integral. For FSK signalling, Montgomery's result is

$$p(e) = 1/2 \; e^{-(S/N)^2},$$

where $(S/N)^2$ is the in-band signal-to-noise ratio. These equations are plotted in Fig. 2.3 together with a curve of measured error rate for the PSK breadboard operating at 51 bits per second. It can be seen that for error rates of less than 1% the predicted advantage for PSK is about 1.5 db, and that this theoretical result has been achieved experimentally for signal-to-

noise ratios above 2.5 db. The reason for the departure from theory at low signal-to-noise ratios was not determined; whatever the cause, the results of the calibration check given in Table 2.1 also indicate poor performance at low SNR, and it is probable that equipment non-linearity and calibration difficulties would account for the discrepancy in both cases.

The results of the above test showed that PSK signalling offered some advantages under ideal conditions with a noise free reference. The next step is to perform service tests with a derived reference, and for this purpose PSK signalling and time base components which can be retro-fitted on the existing data sets have been made.

2.3 PSK Signalling Set

2.3.1 Components

To convert the data sets from FSK to PSK signalling, additional circuit packages for the ground and one airborne set have been provided. A mode change switch is included to permit rapid change-over. The new circuits for the ground unit include the following:

- one crystal controlled time base for operation at 60 bits per second with an oscillator frequency of 1620 cps;
- one PSK keyer for operation with a 1620 cps

sub-carrier;

- one PSK converter for operation with a 1620 cps sub-carrier;
- one 250-watt power amplifier, used to provide synchronous power derived from the crystal controlled time base to the input tape machines.

For the Airborne unit, the additional circuits include

- one time base for operation at 60 bits per second, incorporating a high-Q crystal filter operating at 3240 cps;
- one PSK keyer for operation with a 1620 cps sub-carrier.

2.3.2 PSK Keyer

Fig. 2.4 illustrates the operation of the PSK keyer. A transformer converts the 1620 cps sub-carrier into two sinusoidal components having a phase difference of π radians. Two transistor switches couple these sinusoids alternately to the input of an amplifier circuit under the control of a binary coded message. The amplifier output consists of a 1620 cps sinusoidal wave which changes phase by π radians for each reversal of the binary code, thus fulfilling the requirement for $\frac{\pi}{2}$ phase modulation.

2.3.3 PSK Converter

Fig. 2.5 shows a functional block diagram of the PSK converter. The circuit is essentially the same as that of the breadboard converter described in 2.2.1 with the addition of a peak clipper and a derived reference for the phase detector. The peak clipper removes large impulses from the signal before it is applied to the quenched filter and has a self-adjusting clip level which follows the average signal amplitude.

The quenched filter uses an LC tuned circuit operating in a positive feedback loop to obtain Q multiplication[27]. Quenching is obtained by applying very heavy damping momentarily across the tuned circuit. The filter output passes via a linear amplifier to the phase detector.

The phase detector reference is obtained from the signal as follows. Part of the output from the quenched filter passes to a full-wave rectifier circuit, providing a double frequency component which is independent of phase reversal on the input signal. Since the amplitude of this component will be zero at the start of every bit due to the operation of the quenched filter, a high-Q filter is used to separate it out and ring over the intervals when the quenched filter output is small. The reference signal is finally obtained by limiting and division by two.

The reference signal may have a phase ambiguity of π radians, so that the polarity of the phase detector output may be reversed. A circuit has been added in the airborne set which detects reversed polarity in the block sync pattern and corrects the phase of the PSK converter reference. In the ground set, no provision for phase correction has been made because the air-to-ground message format as originally constituted does not provide a code pattern suitable for detecting reversed polarity signals. Hence, the output tape from the ground set may require manual interpretation; the airborne set will interpret the received signal automatically.

Timing for the airborne set is derived from the high-Q filter in the converter reference circuit. During air-to-ground transmission, the received signal and hence the excitation to this filter disappears, so that timing depends on ringing of the stored energy. The necessary accuracy and high-Q have been obtained by means of a quartz crystal operating in an active filter circuit. In laboratory tests, satisfactory timing has been derived from this circuit for up to ten seconds after removal of the excitation.

The output of the phase detector, a sequence of positive and negative going pulses with waveshapes roughly resembling a sawtooth, feeds a decision circuit which

produces a regenerated message output if the magnitude of the phase detector output exceeds a null reference voltage. A null output occurs when the magnitude of the input to the decision circuit is less than the null reference voltage.

2.3.4 Timing and Synchronization

Exact synchronization between the ground set and its input tape reader was obtained originally through the use of 60 cycle power for operating the tape reader's synchronous motor and the time base of the data set. With the introduction of a crystal-controlled time base for PSK signalling, this synchronizing method no longer applies. The modified approach is shown in Fig. 2.6. Here, a stable crystal-controlled 60 cps sinusoid is derived from the sub-carrier oscillator by divider and filter circuits. A 250-watt amplifier raises the power to a suitable level for operating the input tape reader and existing timing pulse generator.

The method of synchronizing clock pulses in the airborne set is illustrated by Fig. 2.7. It will be remembered that the converter develops a steady tone at twice the sub-carrier frequency for use as a reference signal on the phase detector and that the sub-carrier frequency is exactly 27 times the pulse rate.

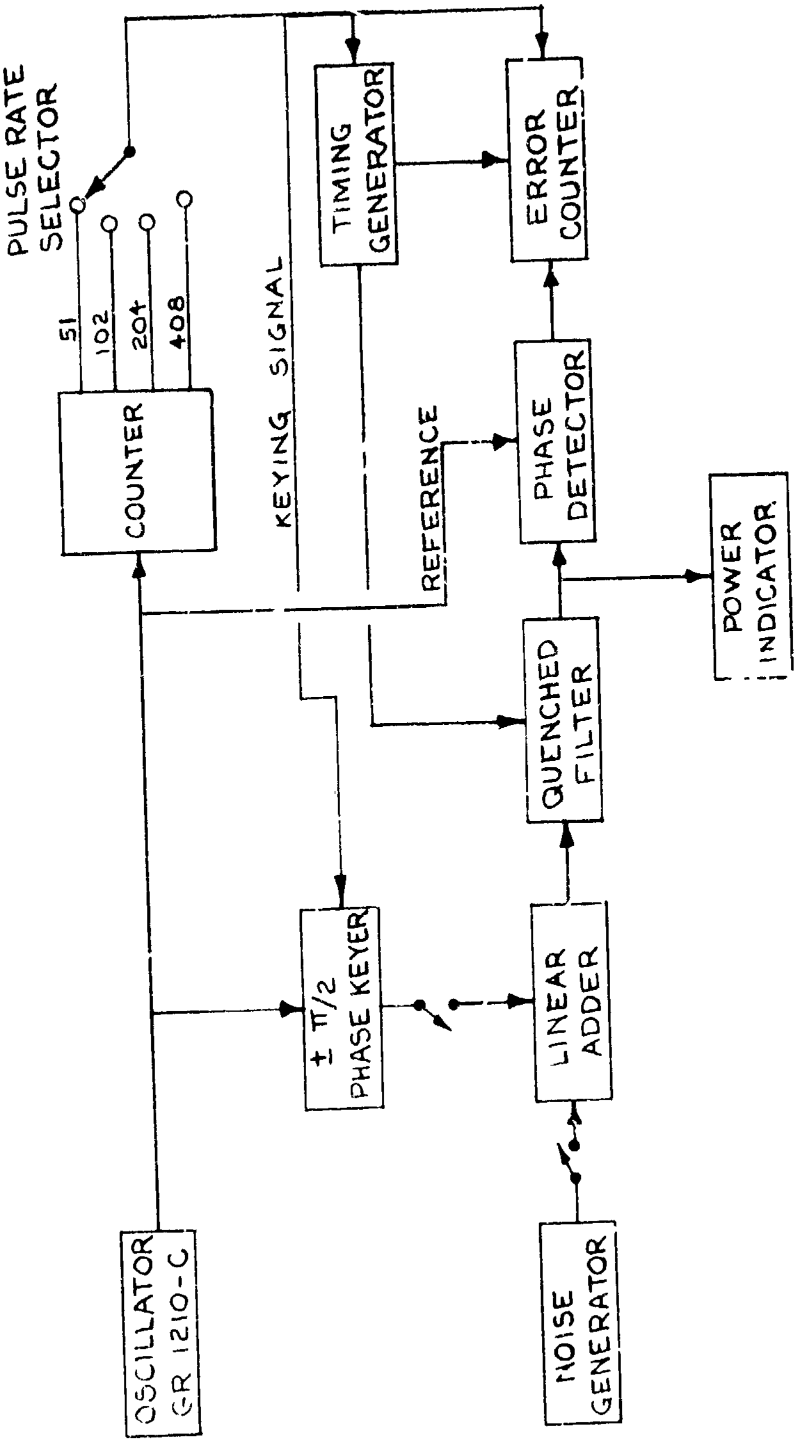
The pulse rate can, therefore, be established at the PSK receiver by dividing the double frequency reference tone in the converter by 54, and complete synchronization can be attained by adjustment of the phasing in the count-down circuit.

In the circuit of Fig. 2.7, a five-stage binary counter circuit is provided with two feedback loops which constrain the counter to re-cycle for either 26 or 28 input pulses, depending on whether the "early" or "late" reset loop is open. The "early-late" signal is derived by comparing clock pulses derived from the counter with the received signal. If the clock pulse is late with respect to the signal, the counter re-cycles for each 26 input pulses until it is early; if the clock pulse is early, the counter re-cycles for each 28 input pulses until it is late.

Operation of the "early-late" control circuit depends on a narrow filter centered on the sub-carrier frequency and quenched at twice the bit rate. Fig. 2.8 illustrates the behavior. During a steady marking signal, the filter reaches about the same energy level between each quenching pulse, but a change from mark to space (or vice versa) results in a reduced energy level in one integral due to the signal reversal when

the quench pulses are not timed correctly. By noting whether the first or second integral within a bit interval as defined by the receiver clock is larger, it is possible to determine whether the receiver timing pulses are early or late with respect to the signal phase reversal. Early and late gates are formed at the end of each bit by storing the value of the first integral and subtracting it from the value of the second. A positive result produces an "early" pulse; a negative result, a "late" pulse.

It will be noted that there is no provision for recycling on 27 pulses, so that the circuit will tend to hunt around the condition of zero phase error. The magnitude of error should average about one sixty-fourth of a bit on a signal consisting entirely of reversals; and on a typical signal, about five sixty-fourths of a bit. A prolonged series of marks or spaces may result in synchronizing errors if the "early-late" control circuit is unbalanced. This difficulty might be overcome by incorporating a threshold in the control circuit and providing an output to control a third feedback loop on the counter for causing recycling on 27 pulses.



PSK BREADBOARD

FIG. 2.1

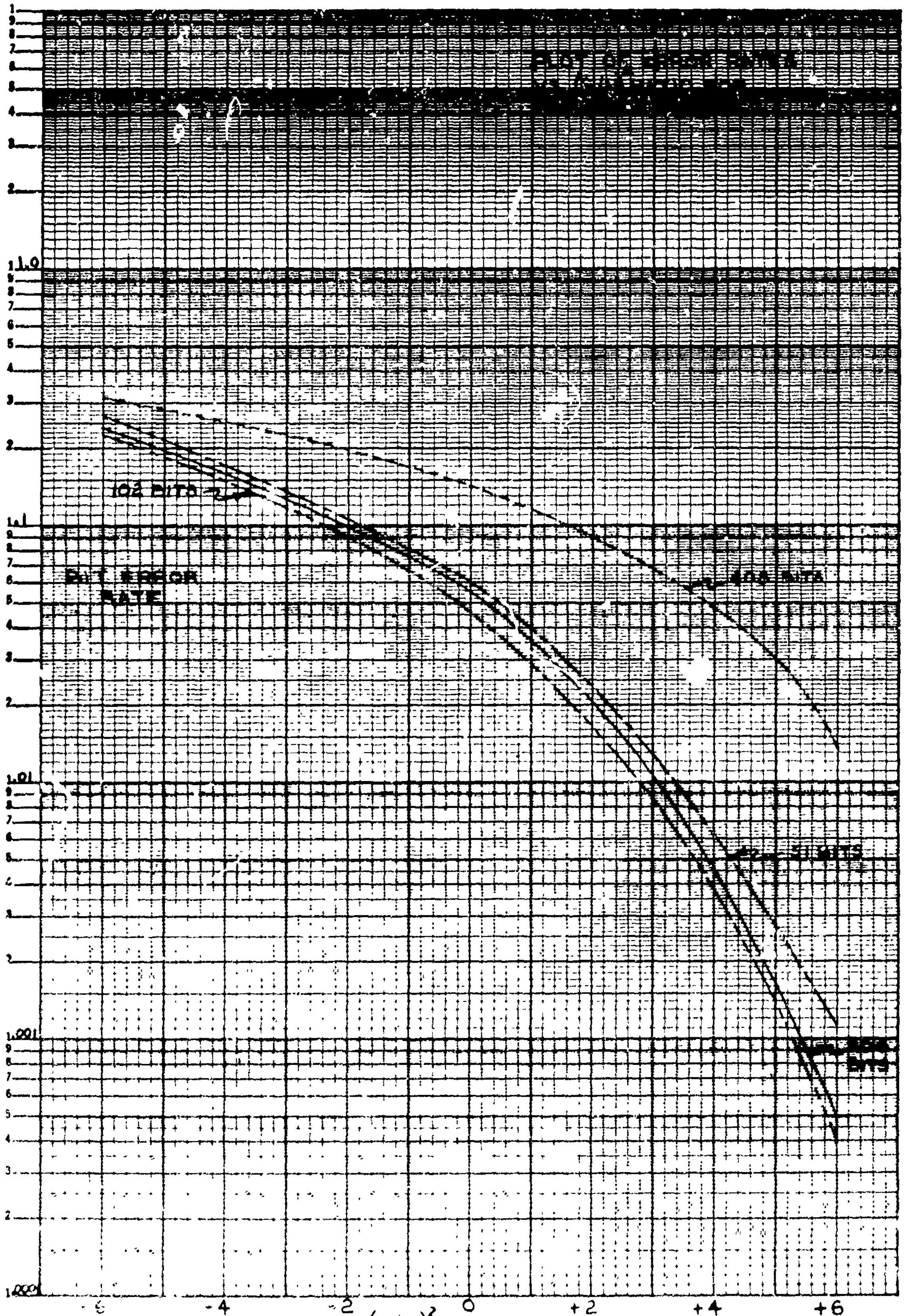
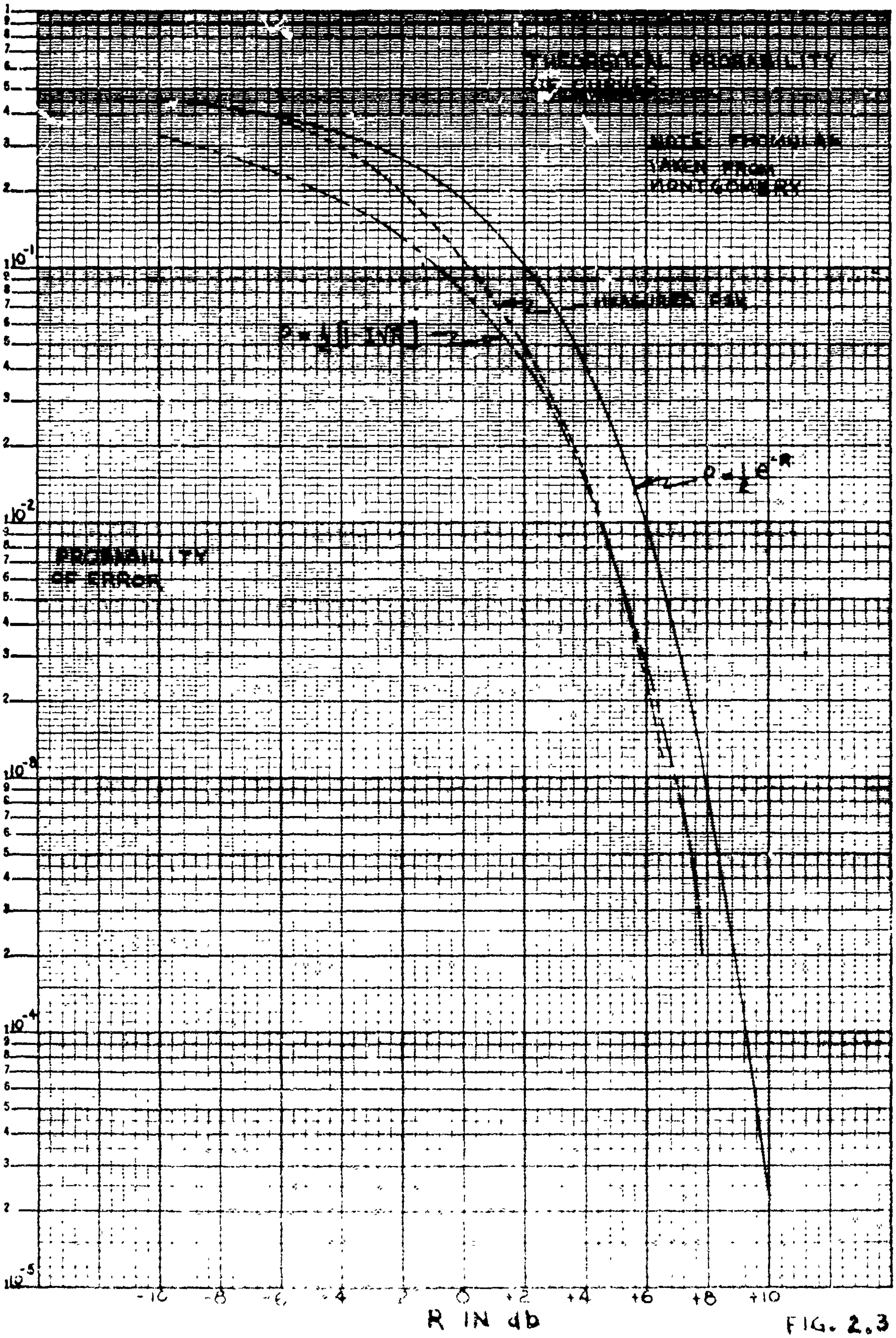
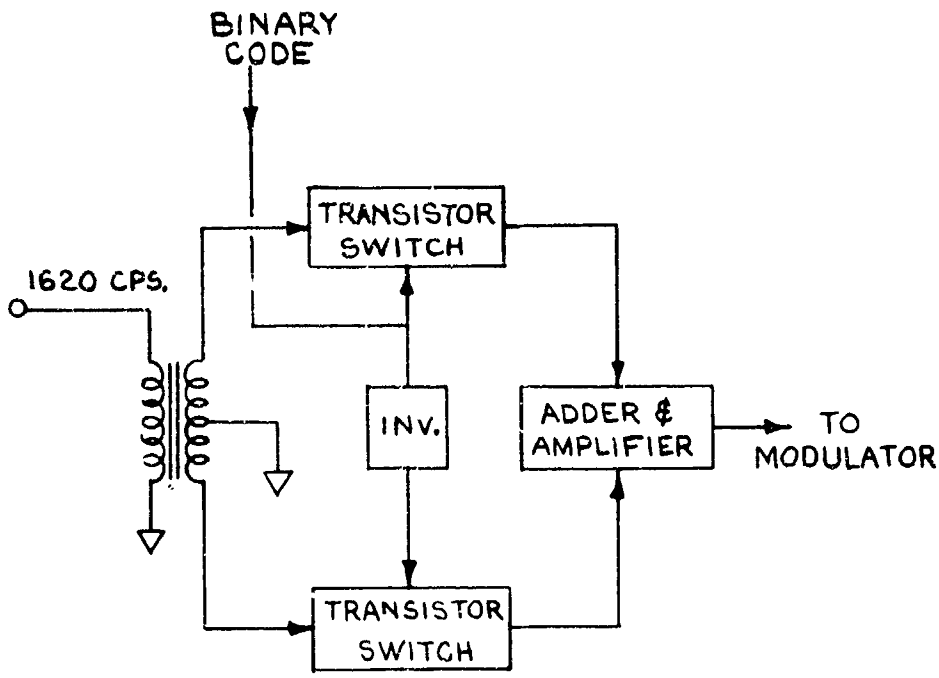

BIT ERROR RATE
102 BITS
408 BITS
51 BITS
1.0
.1
.01
.001
-6
-4
-2
0
+2
+4
+6
$(S/N)^2$ IN db

FIG. 2.2.



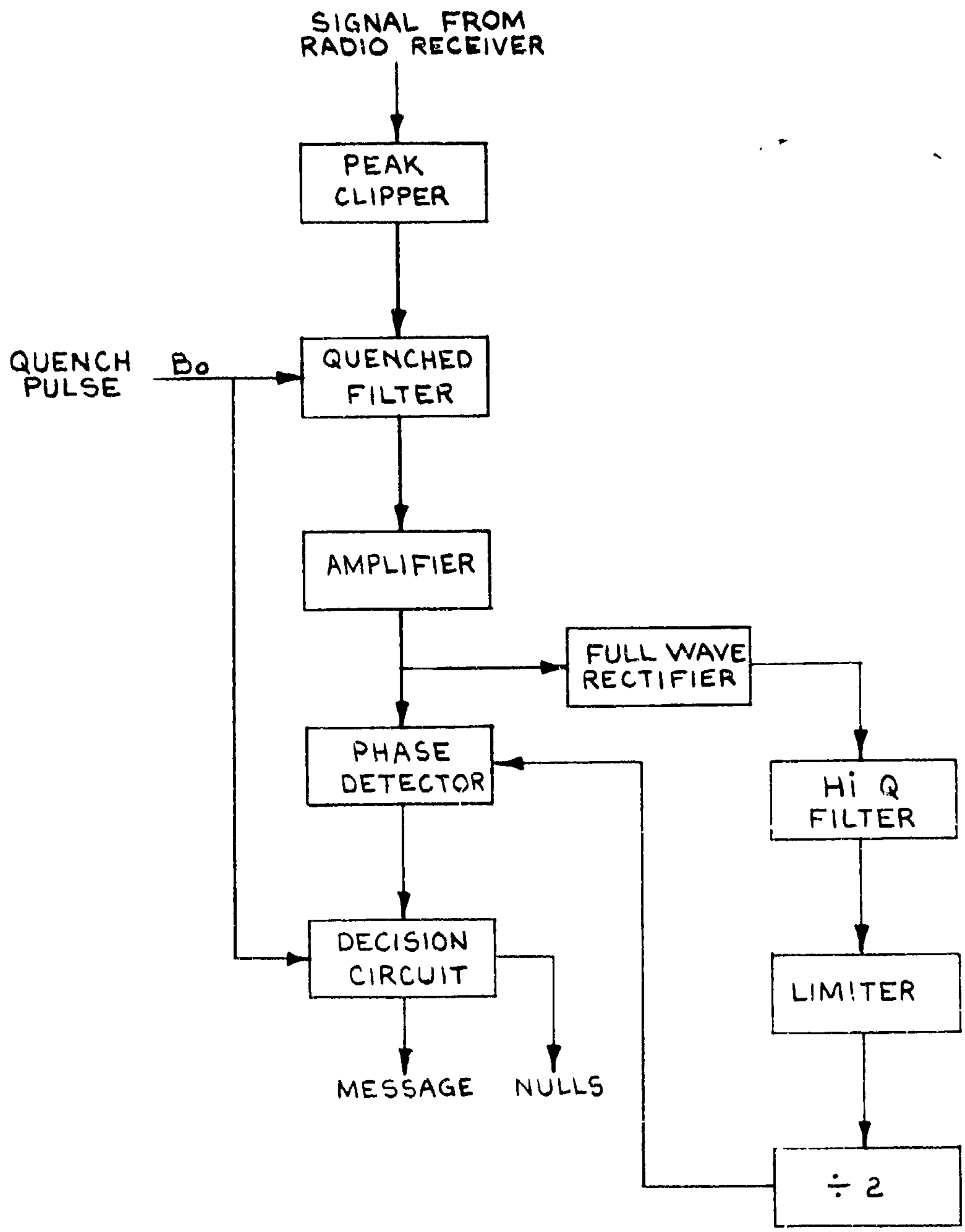
THEORETICAL PROBABILITY OF CURVES
NOTE: FORMULAS TAKEN FROM MONTGOMERY
MEASURED FSK
PROBABILITY OF ERROR
$P = \frac{1}{2}e^{-R}$
R IN db

FIG. 2.3



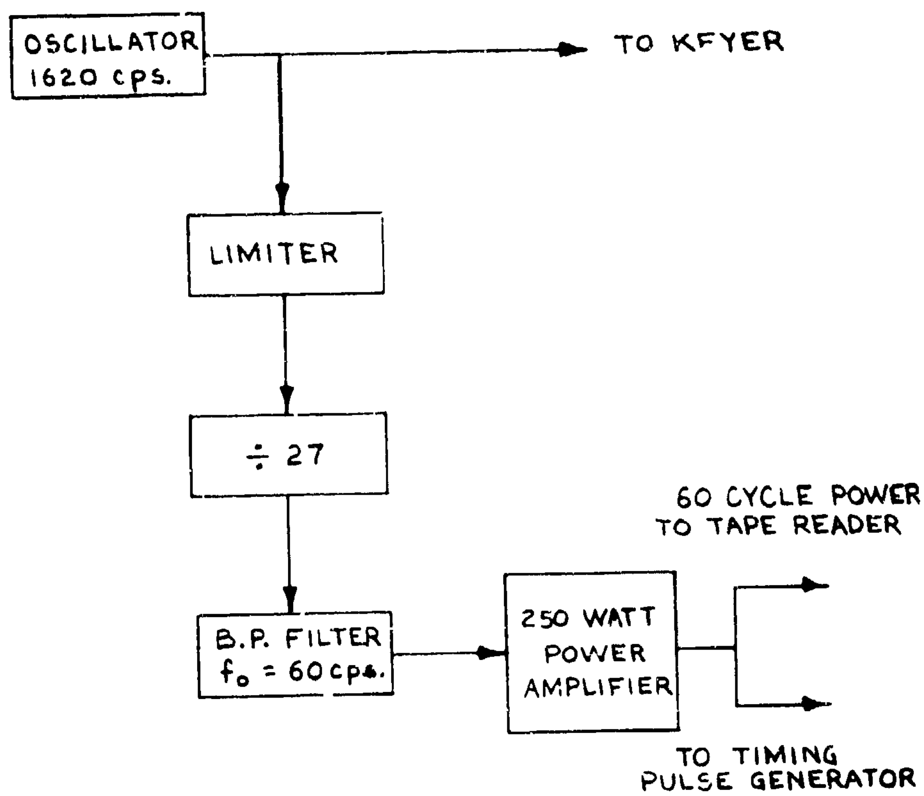
PSK KEYER

FIG. 2.4



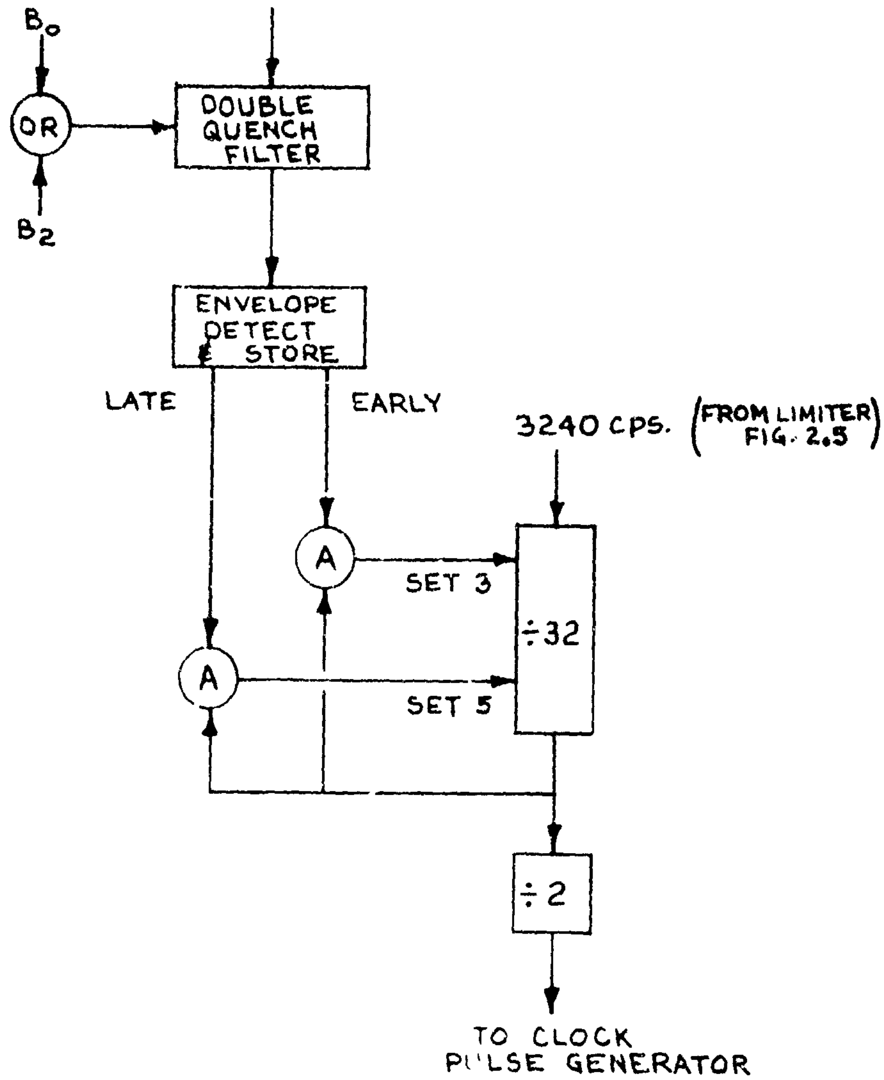
PSK CONVERTER

FIG. 2.5



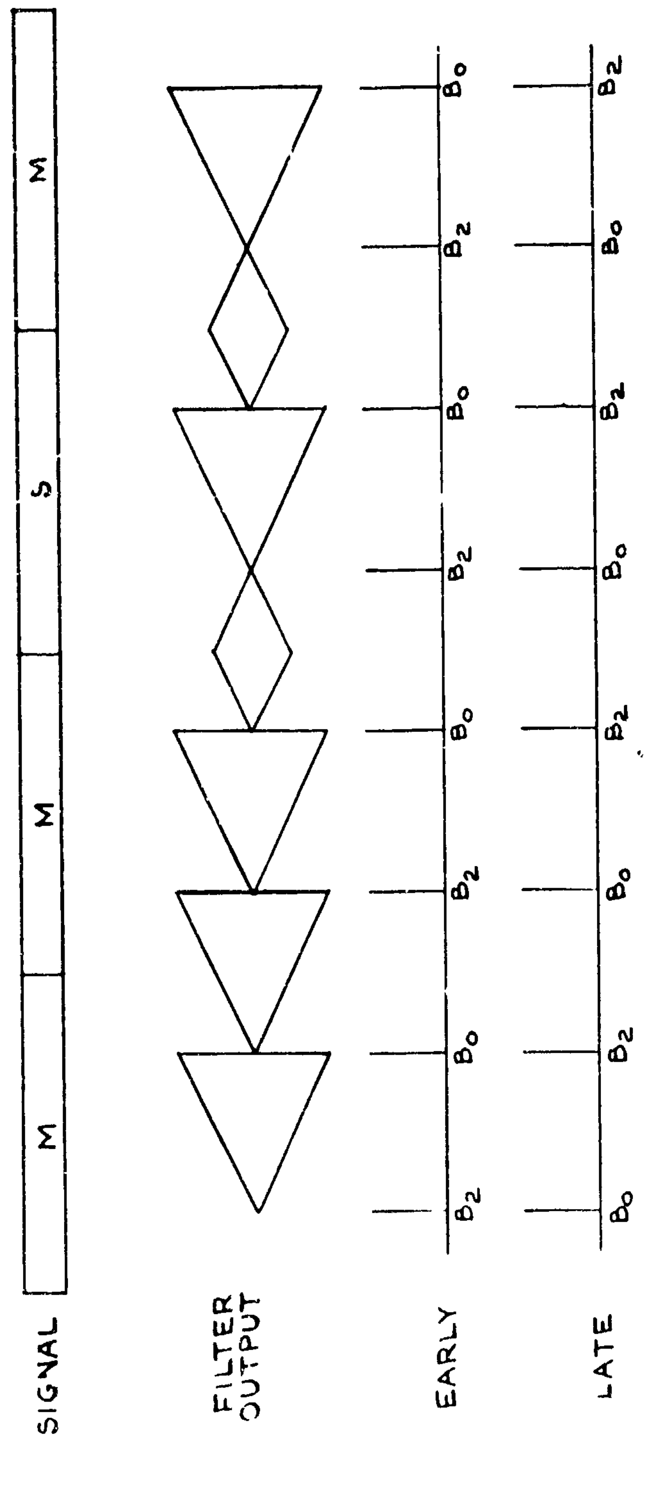
GROUND SET SYNCHRONIZATION

FIG. 2.6



AIRBORNE CLOCK PULSE SYNCHRONIZER

FIG. 2.7



DOUBLE QUENCH FI TER

FIG. 2.8

SECTION 3.0 TEST FACILITY (PART I)

3.1 Introduction

In making performance measurements on any signalling scheme it is necessary to generate the signals, pass them to a receiver via a path in which the signal is disturbed in some way and then determine how well the reciever interprets them. The following section presents a brief functional description of a test facility designed to perform these functions for FSK sub-carrier signals disturbed by random noise. A detailed description of circuits and operating procedures is given in Section 3, Part II of this report.

3.2 Test Requirements

The test objective was to measure the operating characteristics of a group of six converters, each designed for operation at a specific pulse length in the range from 1 to 30 milliseconds, in the presence of random noise. At the same time, the effectiveness of an error correction method based on parity check and null zone reception was to be evaluated. It was accordingly necessary to generate a suitable parity-checked code sequence at each of six signalling speeds. This code sequence modulated an FSK oscillator, the output of which was mixed with random noise and fed to an FSK converter. Signal-to-noise ratio was measured after the band selection filter in the converter. The output of the converter was examined in a group of logic circuits and the significant results totaled on 8 counters.

3.3 Methods

A functional diagram of the test facility is shown in Fig. 3.8. Beginning with a set of five manually operated encoding switches, the message passes in parallel into a transmitter shift register where it is converted to serial form at the appropriate speed. As it flows through this register, an odd parity injector circuit adds a sixth bit so that each code character contains an odd number of marks. The register output keys a frequency shift oscillator directly, and the oscillator output is combined with noise in a linear adder to obtain a test signal for presentation to the frequency shift converter.

The converter circuit includes a pick-off following its input band pass filter which allows signal and noise to be measured in the same bandwidth for determining SNR. Section 4, Part I, presents a functional description of the converter itself; a detailed circuit description appears in Section 4, Part II.

The converter output feeds a synchronously sampled binary decision circuit and a null detection circuit. The binary decision circuit feeds receiving shift register number one, which accordingly accumulates the unmodified output of the converter, and also a bit reversal logic circuit which produces an inverted output for all bit decisions on which null detection occurs. Receiving shift register number two accordingly accumulates the converter output modified by a reversal of all those decisions in which a null is detected. Both registers

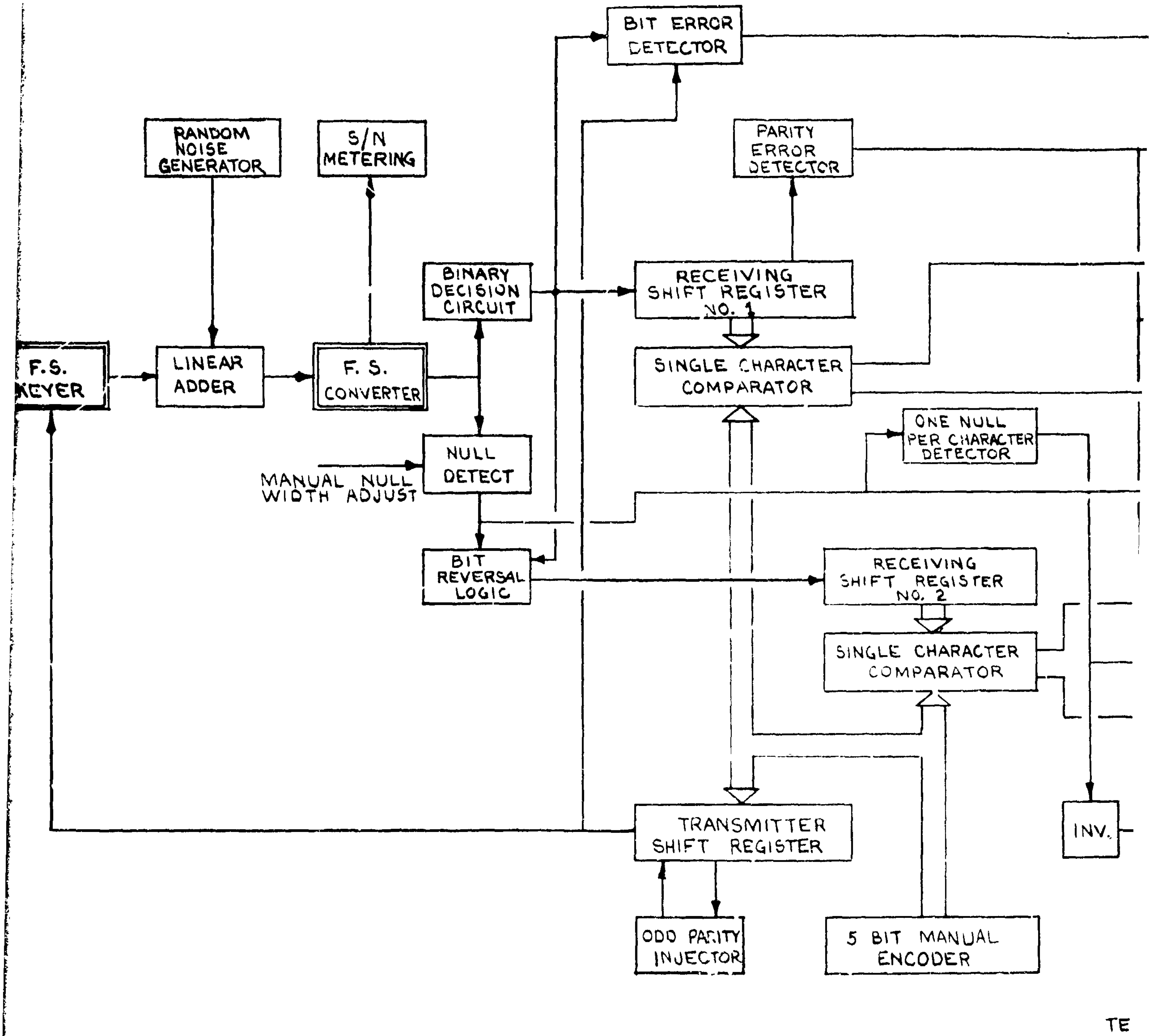
feed parallel comparator circuits which compare the register contents with the manually encoded signal, so that by examining the comparator outputs at the end of each received character it is possible to determine whether the received character is correct in register one or, if incorrect, whether the modified received character in register two is correct.

The bit error detector compares the output of the binary decision circuit with the transmitted code sequence and operates the bit error counter (MC-2) when a discrepancy occurs. At the same time a total null counter (MC-3) registers all outputs from the null decision circuit. All characters received correctly are totaled on the "correct character counter" (MC-1); all those with a parity error in the "parity error counter" (MC-4); and all those containing undetected error (i.e., error but no parity failure) in the "character error counter" (MC-5). Operation of the character error counter MC-5 is blocked for all received characters in which parity fails. Consequently, the sum of the numbers in Counters MC-1, 4, and 5 equals the number of characters transmitted.

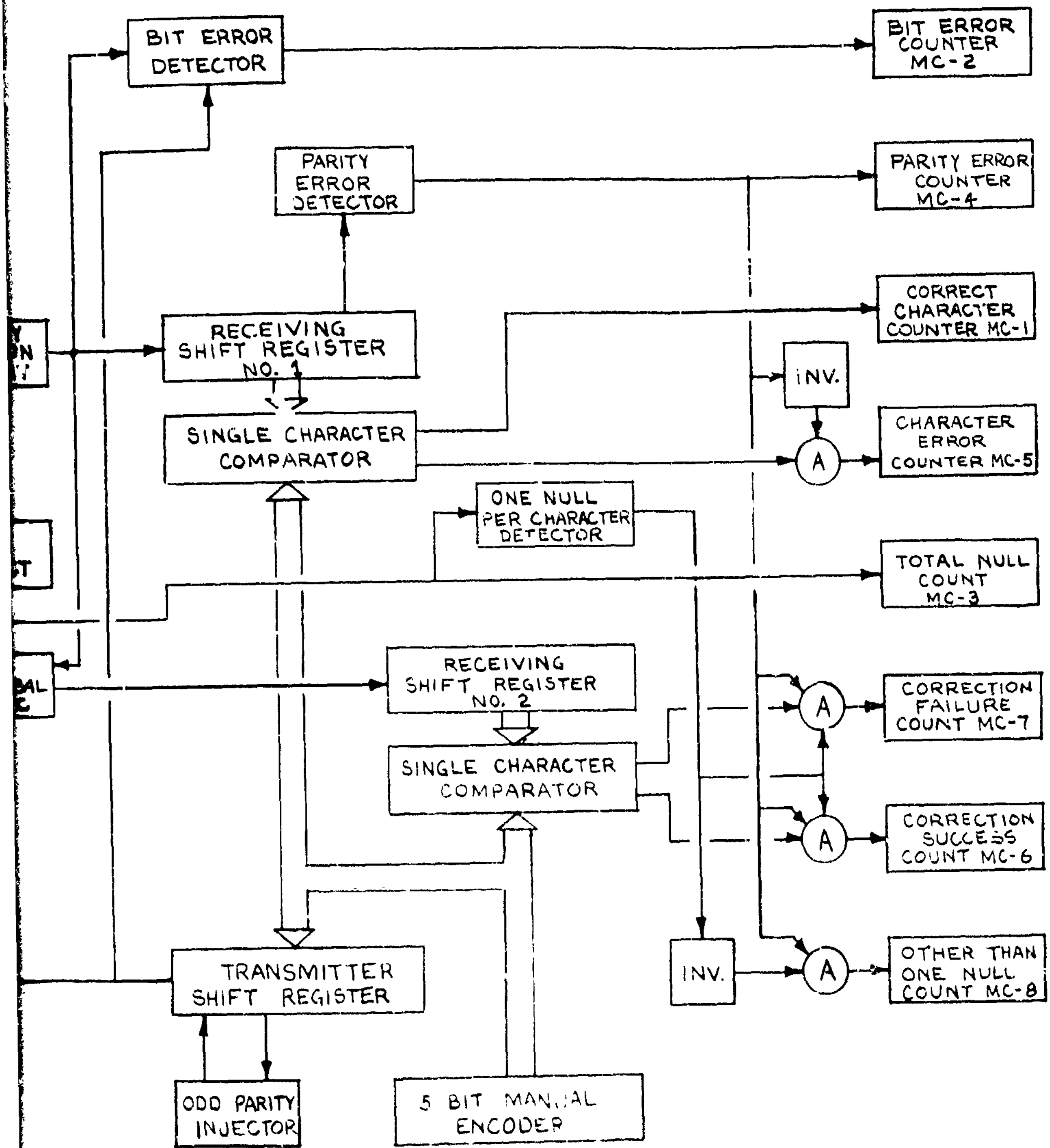
Of all the characters received incorrectly, an attempt can be made to correct only those on which parity failure occurs, since the parity check is the only criterion available to the receiver in judging the validity of the received message. Accordingly, the parity errors totaled by MC-4 may be analyzed further to determine the number of successful corrections (MC-6), the number of correction failures (MC-7), and the number of times when

no correction is possible (MC-8). The criterion which is used in determining whether an attempt should be made to correct a character on which parity failure occurs is the existence of a single null within the character. The number of attempts at correction is accordingly the sum of the indications on counters MC-6 and MC-7; the number of times that no correction is attempted is recorded on counter MC-8; and the sum of the numbers in all three counters (MC-6, MC-7, and MC-8) will equal the number of parity errors as toteled in MC-4.

The test facility includes plug-in units for changing keyer and converter circuits to accomodate each pulse length. Operating speed is selected by a rotary switch; and null width, by a potentiometer adjustment.


BIT ERROR DETECTOR
RANDOM NOISE GENERATOR
S/N METERING
PARITY ERROR DETECTOR
BINARY DECISION CIRCUIT
RECEIVING SHIFT REGISTER NO. 1
F.S. KEYER
LINEAR ADDER
F. S. CONVERTER
SINGLE CHARACTER COMPARATOR
ONE NULL PER CHARACTER DETECTOR
NULL DETECT
MANUAL NULL WIDTH ADJUST
BIT REVERSAL LOGIC
RECEIVING SHIFT REGISTER NO. 2
SINGLE CHARACTER COMPARATOR
TRANSMITTER SHIFT REGISTER
INV.
ODD PARITY INJECTOR
5 BIT MANUAL ENCODER
TE

A



TEST FACILITY - FUNCTIONAL DIAGRAM

FIG. 3.8

SECTION 4.0 CONVERTER DESIGN (PART I)

4.1 Introduction

The FSK converter design used in all the work reported herein is based on the phase shift method as commonly exemplified by the Foster-Seeley discriminator[30]. In a narrow-band device operating at audio frequencies the circuit details, presented in Section 4, Part II, of this report, are quite different and comparison with the Foster-Seeley circuit appears rather far-fetched on the surface. The following section presents the operating principles of a phase shift discriminator and discusses the filter and bandwidth criteria briefly.

4.2 Operating Principle

Fig. 4.13 shows the functional block diagram of a phase shift discriminator (or converter, as it is more commonly called in narrow-band sub-carrier multiplex systems). The first element is a bandpass filter designed to maximize the SNR presented to the limiter circuits. The limiter removes amplitude modulation from the signal and essentially removes any requirement for automatic gain control since it will operate properly over a very large dynamic range.

From the limiter output the signal splits into a direct path to the phase detector and a second path to the reference input on the phase detector via a phase shift network and second limiter. Phase differences between the two signals presented to the phase detector result in an output from which the sub-carrier com-

ponents are removed by a low-pass filter, leaving a DC component which essentially duplicates the modulating waveform on the FSK signal at the converter input. The final step is to square up the output from the low-pass filter in the slicer or level decision circuit.

4.3 Design Factors

Critical factors in matching the converter to the signal are the bandwidth and center frequency of the input filter and the cut-off frequency of the output low pass filter. Goldman[11] shows that the post-filter SNR is nearly maximized for DC pulses when the filter bandwidth is $3/4T$, where T is the pulse width. It follows that the optimum bandwidth for maximizing the post-filter SNR of a pulsed sinusoid is $1.5/T$.

Regarding the FSK signal as a superposition of two pulsed sinusoids, one can see that the input filter might consist of two bandpass filters, one centered at each of the keyed frequencies, and having a total noise bandwidth of $3/T$.

Fortunately, it is not necessary to employ two filters when the separation between the two frequencies is small - say in the neighborhood of $3/4T$. McCoy[31] shows that the output frequency of a bandpass filter excited by a frequency step follows a time function similar to that of the amplitude response of the same filter excited by a pulsed sinusoid and that the behavior is not vastly different if the frequency step begins near the nominal edge of the passband (i.e., the 3 db point) and moves toward

the center. Thus it appears that a nominal filter bandwidth of 1.5/T will accomodate an FSK signal having total shift of 3/4T, and the results presented in Section 5 bear out this deduction.

An important factor bearing on the sensitivity and linearity of the converter is the behavior of the phase shift network. Given a linear phase detector, the phase shift network should introduce a shift in phase of the reference signal which is a linear function of frequency. Since this network normally consists of a bandpass filter in some form the phase shift is not a linear function of frequency, although when the total frequency shift is restricted to a small portion of the bandwidth of the phase shift network the approximation is very good. Such a restriction, however, results in poor sensitivity, since the total phase shift that can be generated is small. The converters employed in the tests reported in Section 5 used a bandpass filter similar in all respects to the input filter as a phase shift network. The linearity and sensitivity achieved is shown in Fig. 4.6.

To obtain zero output at the center frequency, the two inputs to the phase detector must be in quadrature, and the detector itself must be accurately balanced. Use of a two section shift network results in zero phase difference at the reference frequency, and it is accordingly necessary to include a fixed 90° phase shift in the phase shift network to obtain a polar output from the discriminator.

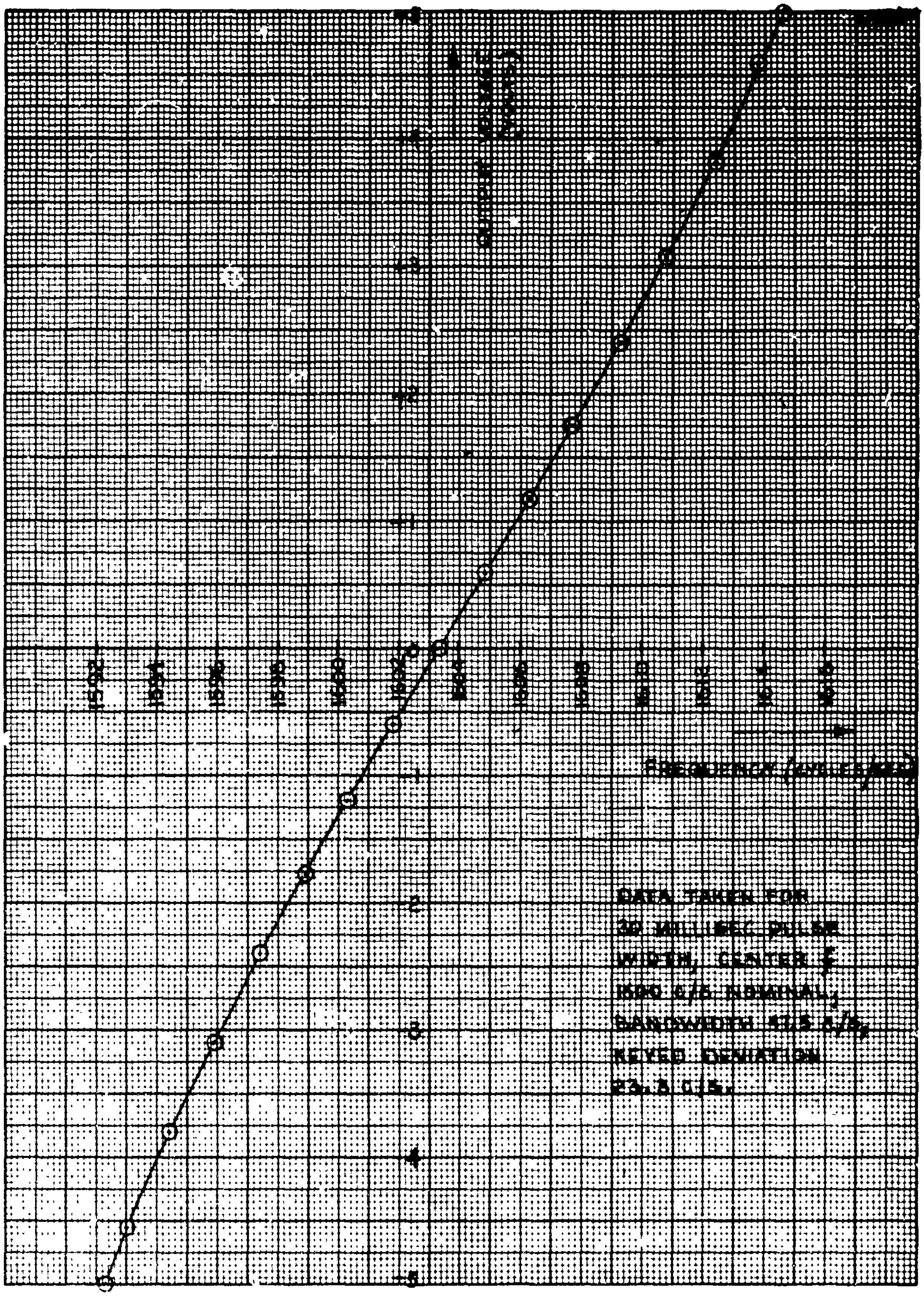
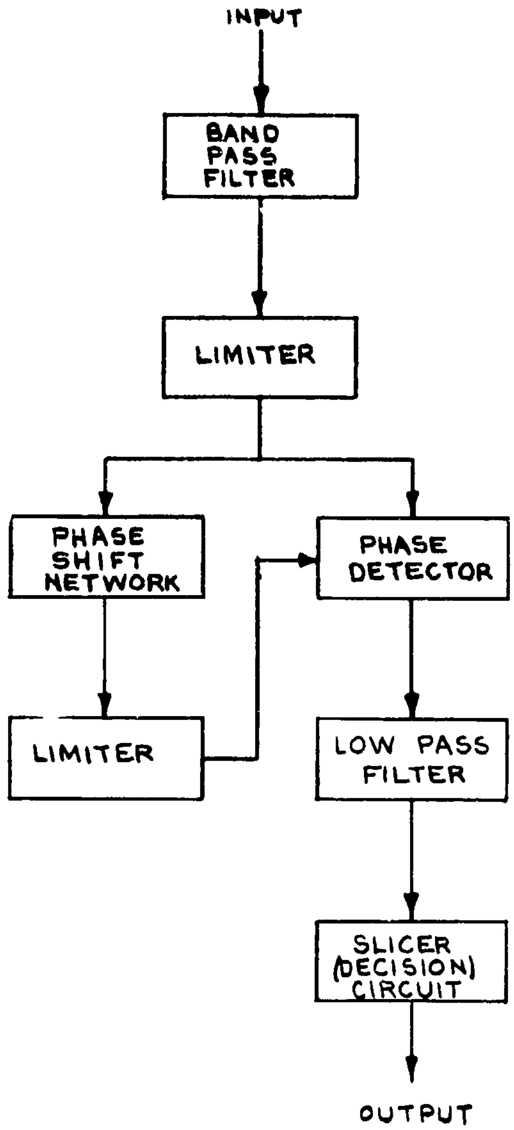


FIG. 4.6



CONVERTER FUNCTIONAL DIAGRAM

FIG 4.13

SECTION 5.0 TEST RESULTS

5.1 BER vs. SNR

5.1.1 Introduction

The criterion for optimizing the FSK converter for each pulse length was the performance of the converter under conditions of random noise. In order to carry out the experiments necessary for this evaluation, a careful study of the desired objective was made. Investigation of the converter as a function of each selected parameter was limited only by the time available to set up and perform the required tests.

5.1.2 Theory

Montgomery[3] indicates that the probability of error is

$$P\ (e) = 1/2e^{-(S/N)^2} \qquad (1)$$

for an FSK system in which the signal to noise power ratio $(S/N)^2$ is measured at the input to the detecting device. A plot of this equation yields a curve which applies to systems wherein the received bandwidth is equal to twice the bandwidth of the equipment following the detecting device based on an optimum transmission rate. Law [10] substantiates these statements for an FM system which uses limiting and center sampling.

Optimizing the bandpass for maximum transmission speed is carried out by Goldman[11] in the bandpass and low pass situations where an approximation in the bandpass situa-

tion gives

$$BW = 3/2 \times 1/t_o, \qquad (2)$$

where BW is the bandwidth in c/s

and t_o is the pulse length in seconds.

Noise power N^2 is a function of bandwidth BW and noise power density N_o^2 per unit of bandwidth so that

$$N^2 = N_o^2 \quad \times \quad BW. \qquad (3)$$

If Goldman's relationship for bandwidth and pulse length (2) is assumed, it follows by substitution into (3) that

$$N^2 = N_o^2 \times 3/2 \times 1/t_o. \qquad (4)$$

Substitution into Montgomery's equation (1) yields

$$P\ (e) = 1/2e^{-\frac{2}{3}\frac{S^2 t_o}{N_o^2}}. \qquad (5)$$

Now let $$P\ (e) = 1/2e^{-kt_o}, \qquad (6)$$

where $k = 2/3\ (S/N_o)^2$. (7)

From this last equation it follows that at $t_o = 0$ the BER (bit error rate) is 5×10^{-1}. This pair of coordinates is used as one point in plotting curves for k. Fig. 5.1 shows a plot of BER vs. Pulse Length for four values of k. These plotted values of k are calculated using the optimum bandwidth factor of Goldman (equation 2) and Kotel'nikov[12], the latter for a passive space system, for constant noise power den-

sities derived from SNR = 2 db and SNR = 3.5 db in a 150 c/s bandwidth.

These curves show that a continuous improvement in error rate can be attained as pulse length increases when signal and noise power density are held constant. In addition, they show how optimized converters would perform in the same noise density. Curves of equations (6) and (1) will be used for converter evaluation.

The first objective of this study is to derive and present a usable means for comparison and design of FSK systems. The second is to show that an FSK converter designed according to equation (2) will satisfy the predicted BER of equation (1). The third objective is to show that such systems are optimized for the design pulse length. The final objective is to show that BER vs. SNR for optimized systems is independent of pulse length when SNR is measured in the optimum bandwidth.

On the basis of equations (1) and (2) the following theory was evolved to show the relative capability of each converter. The improvement in converter operation to be gained by changing the pulse length is

$$M = 10 \log \frac{PL_2}{PL_1}, \qquad (8)$$

where M = improvement in db of SNR for a constant $(S/N_0)^2$ ratio,

PL_2 = new pulse length, and

PL_1 = reference pulse length.

The bandwidth required for a pulse length of 1 ms is 1500 c/s by equation (2). Let 100 equal the signal power required to produce a BER of 1×10^{-1} in this bandwidth. The following signal powers will then be necessary to produce the same BER in a lesser bandwidth according to equation (5):

Pulse Length	Signal Power
1 ms	100
2	50
5	20
10	10
20	5
30	3.33

This theory assumes a constant noise power density in power units/cycle (input SNR is constant for all systems). It follows that each BER will give an appropriate signal power for any pulse length.

1×10^{-1} BER occurs at SNR = 2.1 db power ratio or a numerical power ratio of 1.62. Therefore if $S^2 = 100$ power units, then $N^2 = 100/1.62$ power units.

Normalizing the noise power to a unit bandwidth yields

$$N_o^2 = N^2/BW, \tag{9}$$

where N^2 = noise power density per unit bandwidth,

N^2 = total noise power in the bandwidth in power units, and

BW = bandwidth.

Then by substitution

$$N_o^2 = \frac{61.7}{1500} = .0412 \text{ power units/cycle.} \tag{10}$$

Fig. 5.2 is a plot of BER for signal power in decibels above one signal power unit. Fig. 5.3 presents the same result based on actual signal power units. These curves are based on equations (1) and (2) and the noise power density found in (10). Fig. 5.4 and Fig. 5.5 show this theory plotted in terms of signal power units and pulse length where BER is a constant.

Consider the use of this information in a practical example. Assume an input noise bandwidth of 5 kc/s representing 1 watt of RMS noise power. What signal is required in this bandwidth to operate an optimized FSK converter at a pulse length of 20 ms and a BER of 1 x 10^{-2}? Referring to Fig. 5.1, a signal power of 10.8 db above one signal power unit is required to give 1 x 10^{-2} BER at 20 ms pulse length. Fig. 5.3 gives this value as 12.3 signal power units. Applying equation (2) and solving for t_0 = 20 milliseconds,

$$BW = 75 \text{ c/s.} \tag{11}$$

Using equation (9) the problem noise density is

$$N_o^2 = \frac{1}{5 \times 10^3} \quad \text{or}$$

$$N_o^2 = 2 \times 10^{-4} \text{ watts/cycle.} \tag{12}$$

Using the proportion

$$\frac{S^2 \text{ power units}}{N_o^2 \text{ power units}} = \frac{S^2 \text{ watts}}{N_o^2 \text{ watts}} \tag{13}$$

and substituting the problem values, we have

$$\frac{12.3}{.0412} = \frac{S^2}{2 \times 10^{-4}} . \tag{14}$$

Whence $$S^2 = .0597 \text{ watts,} \tag{15}$$

or an input SNR ratio of

$$(S/N)^2 = 10 \log \frac{.0597}{1} \text{ or} \tag{16}$$

$$(S/N)^2 = -12.24 \text{ db} \tag{17}$$

Improvement or degradation in BER for other pulse lengths can be seen by checking Fig. 5.4 and Fig. 5.5. This is only an illustrative example of the many uses and value of this theory and family of curves. A verification of these hypotheses will be accomplished by proof of the concepts formulated in the objectives of this study.

5.1.3 Implementation

Since the result of tests would be of statistical nature, certain standards were necessary to establish sample sizes at each SNR. These sample sizes were

determined from the equation

$$n = (1/p \quad - 1)\ K/t, \tag{18}$$

where n = sample size,

p = true population error rate,

K = 1.96 for a 95% confidence band,

t = tolerance (5% p) .

An acceptable margin for error in estimating p was taken as a 5% tolerance (maximum difference between sample reading and true population rate) with 95% confidence bands (probability that repeated samples of a given size will result in bands that will include the true population rate for 95% of the bands). For BER's of 1×10^{-4} to 1×10^{-1} sample sizes ranged from 15×10^{6} to 9×10^{4} bits respectively.

Extension of test times to all six operating conditions indicated total time in the order of 175 hours of continuous testing for BER results and 875 hours for null evaluation in addition to set up and SNR measurement time. It was obvious that the time involved was wholly impractical and that a compromise would have to be made. Revised data times were established by using SNR as a mean and dropping extremely high SNR tests. These are reasonable concessions when the standard for evaluation is examined closely.

Equation (1) establishes the basis for the assessment of converter performance in terms of BER (bit error rate) as a function of SNR. Evaluation will be carried out by comparing the SNR required by different systems to produce a standard BER. Variations of 0.1 db represent small degradations of SNR; however, BER changes could be large for this variation. For example, at a BER of 1×10^{-3} a SNR tolerance of $\pm$ 0.1 db represents a BER tolerance of $\pm$ 15%. Therefore, where equation (18) indicates a large sample is necessary in terms of estimating BER, in actuality a much lesser sample is required for a good statistical result expressed in units of SNR.

Since test results will follow a predicted curve, samples in the area of reasonable measurement times will define the trend by comparison to the theoretical result of Montgomery. Simple extrapolation is the only requirement for extending the result. Reduction of test time also facilitates measurement of SNR and minimizes errors caused by variation in equipment operation and temperature changes.

The following table of SNR and sample sizes was calculated based on a 5% tolerance with 95% confidence that measured BER was with $\pm$ 0.1 db of the true population error rate.

SNR	Bit Errors	S/N Voltage Ratio	RMS Voltage Sum (S + N)
2.0 db	2604	1.259	1.608
3.5	1243	1.496	1.80
5.0	799	1.778	2.04
6.5	362	2.113	2.34
8.0	154	2.512	2.70

Measurement of SNR was carried out by an RMS voltage measurement at the output of the first bandpass filter of noise and signal alone. The foregoing table shows the voltage ratio required and the RMS voltage sum produced for N = 1.

In all cases, insertion loss has been disregarded in calculations. Since insertion losses will effect noise and signal equally, SNR will remain unchanged although absolute values will be affected.

5.1.4 Measured Results

Fig. 5.6A through Fig. 5.6F show the measured BER for each pulse length tested vs. SNR along with Montgomery's (1) theoretical prediction for an FSK system. It will be shown that the measured values define a very near optimum state of operating conditions. Initial inspection will reveal, however, that five of the six curves (all except the 30 ms pulse length) fall inside the theoretical result, i.e., BER's are better than expected. Analysis of this phenomenon would seem to indicate the

deviation was caused by measurement inaccuracies, system idiosyncracies, a noise source which has other than random characteristics, design error, or component tolerance, since the measured result appears to surpass Montgomery's theoretical prediction.

As indicated in 5.1.2 optimization can be shown for a given converter design by a series of BER vs. Pulse Length tests. Fig. 5.7 shows the results of such a test performed with the 10 ms (BW = 150 c/s) system at two convenient SNR A pair of tangents have been drawn to the BER curves for K = 168.8 and K = 244.2, representing a constant noise power density based upon SNR = 2 db and SNR = 3.5 db respectively, in the experimental bandwidth.

The point of tangency falls at what shall be hereafter defined as the optimum pulse length. Although smaller pulse lengths produce higher BER's, it is interesting to note that wider than optimum pulse lengths produce BER's which are less than those of equation (1). However, since BER results of the order indicated by the plots of K = 168.8 and K = 244.2 are readily attainable by optimized systems, no support for operation of this converter at slower bit rates can be realized. Indeed, BER is better at 20 ms than

at the optimum pulse length for this system; however it is much poorer for the same signal and transmission is much slower for the same BER than in an optimized system.

The results shown in Fig. 5.7 would seem to explain the phenomenom appearing in the measured curves of BER vs. SNR. These curves indicate that the pulse length used for SNR tests is somewhat longer than required, i.e., the bandpass is wider than necessary. Reference to the bandpass characteristics of the 10 ms filter (Fig. 4.2D, Part II) indicates an actual 3 db passband of 155 c/s, somewhat larger than the design BW of 150 c/s. For this set of parameters this bandwidth would make Goldman's optimized pulse length equal 9.65 ms, which is extremely close to the result realized. In fact, it can be shown that variations in pulse length of ± 0.5 ms represent BW changes of ± 7.5 c/s. Considering the conditions of design and test, the resulting achievement is an excellent representation of an optimized effort.

Although time was insufficient to run a complete optimum analysis for each pulse length, nor was the Facility entirely compatible for such an effort, it is readily seen how each pulse length has resulted in lower BER's than expected. Examination of all band pass character-

istics except the 30 ms set will show a wider passband than called for by design and therefore a lower optimum t_o and lower expected BER for the design pulse length. The 30 ms filters show a narrower bandpass and therefore longer t_o and higher expected BER for this particular test. Indeed, inspection of Fig. 5.6F shows a BER curve which falls almost exactly on the predicted optimum.

Assessing bandwidth by arbitrarily assigning 3 db points to represent rectangular passbands is an assumption which can easily bias the result and produce slight deviations such as those noted in these experiments. Nevertheless, these systems are very near optimum, and it follows that any system can be as readily optimized. Since these systems perform equally well at the same SNR measured in the bandpass, each system performs successively better as the pulse length increases and the noise density N_o^2 remains constant.

Converter design was carried out to realize a wide dynamic range of signal and noise inputs. Since sensitivities of less than 1 mv were attained at all pulse lengths with a total dynamic range of 70 db, it follows that BER vs. SNR tests at various amplitude levels yielded like results.

5.1.5 Conclusions

It has been shown by these experiments that the series of curves produced to provide the information required for evaluation of any FSK system and for design of optimum systems in 5.1.2 are valid. It has also been shown that (1) systems can be readily optimized using equation (2); (2) all optimized systems regardless of pulse length yield the same BER for SNR measured in the bandpass; (3) BER of an optimized system follows equation (1); and (4) improvements in BER can be readily attained and accurately predicted for the same input SNR, i.e., constant N_O, as the pulse length is increased.

5.2 CER vs. SNR

If p is the probability of bit error, then (1-p) is the probability that the bit will be received correctly. The probability that all six bits in a six bit character will be correct is $(1-p)^6$. Therefore, the probability of character error (CER) is

$$P = 1 - (1 - p)^6$$

or approximately 6p as p approaches zero.

This probability has been plotted in Fig. 5.8A through Fig. 5.8F as a theoretical prediction. The dotted curve is the measured experimental result. It will be noted that these curves follow the same trends as discussed in Section 5.1 on BER's. The

random character of the noise component of the SNR is substantiated by the close conformity of these curves to the theoretical curves.

The bias of the various converter parameters was investigated further under pure noise conditions. These measurements of bit error rates reinforced the concept of random noise. Results tabulated showed negligible deviations from 5×10^{-1} bit rates when the converter was properly balanced.

Messages used in the experiments made mark and space equally probable. The tests of pulse length vs. BER for the same system were performed with an alternate mark space transmitted. The transmitted message for all other tests was mark, mark, space, space, mark, space.

5.3 Center Sampling

5.3.1 Introduction

The Test Facility uses the center sampling technique for mark or space decision. A sample pulse shifts the squared signal output from the converter into the receiving register. This register assumes the condition of the signal at the shift time and maintains this condition until the next sample pulse arrives. The objective of this experiment is to evaluate sample times in terms of BER and SNR and to determine the optimum sampling time for a given system.

5.3.2 Theory

The effect of noise on the received FSK signal is a random displacement of the leading and trailing edges of the detected and squared waveform. The distribution of these edges is a function of the SNR. When noise is sufficient to produce a received time error of more than $\pm$ 50% in the mark or space condition (phase change greater than $\pm$ 90°) an error will result by center sampling. Theoretically noise adds equal amounts of phase lead and lag in the received signal when averaged over a sufficiently long interval. In practice, however, inherent distortion produced in the system creates a condition which only approaches the ideal.

5.3.3 Inplementation

In order to evaluate sampling time, a series of tests were performed on the 10 ms pulse length system. Center sampling was taken as the mean (0%) and sampling times were expressed as a positive or negative percentage of the total bit time. These times were varied to the limit of the Facility which was -35% to +15% for this set of parameters. BER measurements were made for SNR for the standard error tabulation (see Section 5.1.3).

5.3.4 Results

Test results were plotted in Fig. 5.16 in the form of

BER vs. sample Pulse Position (%) for each of the six measured SNR. These curves show an interesting result. Degradation in BER as a function of pulse position is much more severe at high SNR than at low SNR. For example, where SNR = 2 db, a degeneration of 2/1 in BER occurs at a sample pulse position of -26%, while at SNR = 9.3 db the same BER deterioration occurs at a pulse position of -11.5%.

The optimum sample position was computed for all SNR by taking an average mean value based on a 25% degradation from the minimum BER for each SNR. These tabulations show that the optimum sample time is + 2.75% for this system. This represents an error when compared with center sampling; however, the error is negligible since BER is virtually the same for both sample positions.

A family of BER vs. SNR curves were computed and expressed as a percent pulse position based on the mean average where BER for the mean was normalized to satisfy Montgomery's equation for the probability of error. These results are shown in Fig. 5.17 where they can be compared in terms of SNR. The effective loss of signal can be determined from these curves for pulse sampling tolerances.

5.3.5 Conclusions

Evaluation of center sampling of the 10 ms signal showed that the error in sampling time from the mean optimum was only 2.75% of the bit time. The results of this evaluation were normalized to an optimum condition and expressed as a percent pulse position based on the mean average. This presentation shows approximately a 0.3 db degradation in SNR for a sample tolerance of ± 10%.

5.4 Null Evaluation

5.4.1 Introduction

The experimental null zone reception system evaluated by the Test Facility is a parallel binary decision device whose secondary (null) output is the basis for correcting the primary (signal) output when parity error occurs. The system has a single null zone. Correction is limited to those characters identified with odd parity bit errors and single null and decisions are based on simultaneous center sampling of the demodulated signal. In this report a single null zone is the type of measuring standard used to determine the occurrence of a null, whereas a single null refers to a situation of one null per character. The following section discusses a series of experiments performed at each pulse length to evaluate this system of null zone detection and to determine what improvement in character error rate (CER) can

be attained with the correction process.

5.4.2 Theory

Single null correction will be defined for these experiments as a process whereby correction is based upon the detection of a null in a single null zone. A null is detected when the signal amplitude falls between preset null zone limits at the time of sampling. The null zone for single null reception[13] is best explained by consideration of a three-level decoder which assigns the received signal (Y) to one of three groups:

$$\begin{array}{ll} \text{Group I} & y \geq +e \\ \text{Group II} & -e < y < +e \\ \text{Group III} & y \leq -e, \end{array}$$

where Group I records as Mark, Group III records as Space, and Group II records as Null, and +e and -e are the null zone limits.

The binary signal approximates a sinusoidal waveform (for alternate mark-space) which is symmetrical about a center reference level e_o (Fig. 5.9A). The null zone is defined by equidistant positive (+e) and negative (-e) levels with respect to e_o (Fig. 5.9C). The objective of this study is to determine an optimum null zone and assess null zone reception on the basis of the improvement in character error rate (CER) which can be attained by a correction system using this process.

There are three requirements for character error correction in the null system. The first requirement is the detection of bit errors. Parity check is used by the Facility for error detection because of its simplicity and adaptability to this study. The transmitted character consists of five information bits and one (odd) parity bit. Correction is based upon the fact that parity errors are predominantly single bit errors, since the probability of one bit error in six is much greater than the combined probability of three and five in six. Using the general formula[10]

$$P_{cbr} = \frac{r!}{b!(r - b)!} P_e \quad (1 - P_e)^{r - b}, \qquad (1)$$

where P_{cbr} is the probability of b elements being wrong in a character of r elements and P_e is the probability of element error,

and expanding for 1 and 3 bit errors and adding the result expressed as a percent of CE produces curve A in Fig. 5.10A. The term for 5 bit errors has been disregarded because it represented less than .01% of the term for 1 bit error and could not be plotted. Therefore, curve A represents all parity errors (PE) as a percent of total character errors (CE). Curve B in this figure shows the probability of one bit error in six and represents all single bit parity errors (SPE). Note that Curve C (SPE/6CE) is 16-2/3% of Curve B and is the

fraction of all single parity errors that would be successfully corrected by a random process.

Expansion of the binomial (1) for bit error probabilities of no error through six errors for a six bit character with a largest bit error probability of 1×10^{-1} will show that the sum of the last three terms (probability of 4 or more bit errors) is 1.27×10^{-3}. Therefore bit errors are virtually accounted for by distribution of 1, 2, and 3 bit errors per character and can be represented by the following equation:

$$BE = SPE - 2UCE - 3(TPE - SPE), \qquad (2)$$

where BE = bit errors,

SPE = single parity errors,

UCE = undetected character errors, and

TPE = total parity errors.

This equation (2) will be used for analysis of null evaluation test results.

Consider a perfect correction system whereby all single parity errors (one bit error per character) are always corrected successfully. Two curves are shown in Fig. 5.10B which represent character error rate (CER) for an optimized system (Curve A) and character error rate (CER') for the same optimized system (Curve B) where all single parity errors are successfully corrected. Since

$$\mathrm{CER} = \frac{\mathrm{CE}}{\mathrm{CH}}, \tag{3}$$

where CE = character errors and

CH = total characters transmitted,

$$\text{then} \quad \mathrm{CER}' = \frac{\mathrm{CE}}{\mathrm{CH}} - \frac{\mathrm{SPE}}{\mathrm{CH}}, \tag{4}$$

where SPE = single parity errors.

The probability of character error is

$$\mathrm{CER} = 1 - (1 - p)^6, \tag{5}$$

where p = BER (bit error rate), and the probability of one bit error per character[10] is

$$\frac{\mathrm{SPE}}{\mathrm{CH}} = 6p\ (1 - p)^5. \tag{6}$$

Whence (4) may be rewritten as

$$\mathrm{CER}' = 1 - (1 - p)^6 - 6p\ (1 - p)^5, \tag{7}$$

which simplifies to

$$\mathrm{CER}' = 1 - (1 - p)^5\ (1 + 5p). \tag{8}$$

Curve B (Fig. 5.10B) shows the theoretical improvement in character error rate which can be attained by the ideal correction of all single parity errors. It will be noted that the improvement is in the order of 3 db. This value will be used as a standard for evaluating the null correction system. Curve C (Fig. 5.10B) shows the improvement in character error rate by random correction of all parity errors (note: random correction of PE will be successful for SPE/6). Improvement by random correction is slight and by inspection of Curve C

is about 0.2 db.

Null detection appears to provide a means of improving the chances of successful correction of the bit in error. When parity check fails one bit in error is assumed; hence one correction is required. The second requirement for correction is therefore a single null (SN). If no nulls or more than one null have occurred, no correction will be attempted. Fig. 5.9D shows this analysis in block form. The third requirement for a successful correction is that the null and bit error be coincident. The three requirements for character error correction can be expressed as follows:

$$\frac{PE}{CE} \times \frac{SN}{PE} \times \frac{CCE}{SN} = \frac{CCE}{CE}, \qquad (9)$$

where PE = parity errors,

CE = character errors,

SN = single null, and

CCE = successfully corrected character errors.

This expression would equal unity for successful correction of all character errors. However, it has already been shown that if SN/PE and CCE/CE are unity the relationship of SPE/CE will determine what fraction of the total errors can be corrected (Curve B of Fig. 5.10A) and that this ideal correction will give a 3 db improve-

ment (Curve B of Fig. 5.10B) of SNR.

Having established the relationship of PE/CE, it remains to maximize the expressions of SN/PE and CCE/CE to provide the largest correction CCE/CE possible. This will be accomplished by a series of BER vs. Null Width tests for several SNR.

5.4.3 Implementation

The filtered output of the detector centered about reference level e_o is squared by the decision circuit, a voltage comparator with reference e_o. Two identical voltage comparators with +e and -e as references perform parallel squaring functions for null circuit logic. These circuits have high sensitivities and switching speeds are of the order of 40 microseconds. In the presence of noise they will produce random frequency square wave outputs. The signal waveform is center sampled directly after squaring; however, the output of the null comparators must be combined before sampling can take place.

Consider two voltage comparator circuits A and B, with reference inputs +e and -e respectively (Fig. 5.9B). Outputs will then be assigned as follows:

Output	e Signal
1	$>$ e reference
0	$<$ e reference

The output for null C = 1 (Fig. 5.9C) will then be a [illegible]

and b = 1 and the equation for null, in switching circuit algebra, is[29]

$$a'b = c \; .$$

If the "a" and "b" outputs are AND gated after inversion of "a" the proper output for null will be obtained (Fig. 5.9B). This output is made a parallel adjunct to the signal output and center sampled.

The sampled null output is used to correct the signal as follows:

Signal	Null	Corrected Signal
Mark	No Null	Mark
Space	No Null	Space
Mark	Null	Space
Space	Null	Mark

This table shows a bit inversion each time a null is detected. In the Test Facility, two comparators check the received signal character by character against the transmitted signal: the first checks the received signal exactly as center sampled after squaring, while the second checks the same message with bit inversions when a simultaneous null occurred.

Detection of parity error is performed by a counter which operates on mark input and is sampled and reset at the end of the sixth bit. A second counter which registers

the nulls occurring in each character is also sampled and reset at the end of the sixth bit, giving an output when only ONE null has been detected. The total of all nulls received is tabulated by a third counter. When parity error and one null have occurred, the corrected signal in the second comparator is checked. The result of the comparison with the transmitted signal will depend on the number of bit errors in the parity error and the probability that the detected null occurred coincident with the bit in error.

The received signal is checked in the first comparator for all conditions other than single null and parity error. If the first comparison is incorrect and parity has not checked, the character will be counted as a message block with parity error and a no null or more than one null condition. If parity checks, the comparison will reveal a correctly received message (comparison good) or an error other than parity (comparison fails). These operations are shown in function block diagram Fig. 5.9D.

The corrected signal is not used when parity checks even though the message may still be in error, and even though each bit having a coincident null is reversed in the second comparator, for this comparison is not used unless only a single null has occurred. Single nulls may occur

at times when parity checks; however, they are not recorded as single nulls in this analysis but are added to the total nulls received.

The null zone width is varied from 10% to 50% of the detected peak to peak 10 volt signal amplitude for these tests. SNR used were the same as those for the BER vs. SNR tests.

5.4.4 Results

Fig. 5.11A through Fig. 5.11F show the results of tests for each of the six pulse lengths for the rate of occurrence of nulls as a percent of the total bits transmitted for all null widths and SNR. These curves show that increasing null width e_w or decreasing SNR increases the number of nulls detected. The marked similarity of the families of null rate curves justifies pooling the null evaluation results for all pulse lengths to produce a more accurate statistical result and a more representative indication of the value of null detection systems.

Fig. 5.12A shows a range of test results for PE/CE for all SNR tested. These maximum and minimum values represent all pulse lengths and null widths since parity error is independent of these parameters. These results show good agreement with the theoretical prediction of Curve A

in Fig. 5.10A. Using the median values of this test the Curve A (PE/CE) on Fig. 5.12B has been drawn. Curve B (SPE/CE) on this figure was derived from equation (2) using measured results. Random correction of these errors would result in Curve C (SPE/6CE) shown on this figure.

It is imperative that null detection be optimized to produce the greatest number of single null characters. Although single nulls occur when parity checks, they are required only when parity fails (see equation 9). Therefore, optimized SN (single null) is shown in the range of test values of SN/PE in Fig. 5.13A to Fig. 5.13E with null width the independent variable. This family of curves shows that an optimum null width can be obtained for each SNR and by inspection of the curves this optimum occurs at approximately a 30% null width.

After parity error has been detected and a single null has occurred, a correction is attempted, i.e., the second comparison is checked for success. Curve A in Fig. 5.14A through Fig. 5.14E shows a median of the range of test results of CCE/SN, i.e., the empirical probability that the null and bit error were coincident (see equation 9). These curves are shown for the five SNR tested. Curve B in Fig. 5.14A through Fig. 5.14E

shows the range of test results for CCE/CE for each null width. This product (9) is a figure of merit for a correction system. Curve C in these figures is a plot of the successfully corrected characters that could be obtained with the existing system if corrections were made at random (independent of null detection) and is derived from Curve C of Fig. 5.12B.

The curves of CCE/CE substantiate the previous statement that a null width near 30% is an optimum adjustment and maximizes the number of characters corrected successfully by the single null correction system. These curves show that as the SNR increases the improvement in error correction success for a null system increases much more rapidly than for a random correction system. This is shown in Fig. 5.14F where three curves are plotted. Curve A is the median of the range of test results for a null width of 30% (optimum) of CCE/CE. Curve B is a presentation of the correction which can be obtained by a random correction process (CCE'/CE). Finally, Curve C can be interpreted as the correction (SPE/CE) for an ideal system where all single parity errors are successfully corrected.

The final evaluation of an error correction system is made by comparison of the corrected character error rate (CER') with non-corrected character error rate (CER)

in terms of SNR. Fig. 5.15A through Fig. 5.15F show the character error rate for each pulse length before and after correction for four SNR and four null widths. Improvements in the order of 0.4 to 0.7 db are noted for all pulse lengths for the optimum null width or 30%. This improvement is small when compared to the ideal correction system (Fig. 5.10B), yet considerably improved over the random correction process.

Although correction efficiency improves more rapidly with SNR for the null system than it does for either the ideal or random process (Fig. 5.14F), correction becomes far less important at high SNR since error rates are already very good without correction. It can therefore be stated that the results of this evaluation show only small improvements in terms of SNR by a single null detection and error correction system.

5.4.5 Conclusions

A summary of Null Evaluation results shows that

1 - Optimum null widths exist for single null systems,

2 - Null rates are independent of pulse lengths,

3 - Parity errors represent a major portion of the total character errors at usable SNR,

4 - Parity errors are primarily single bit parity errors,

5 - A single null system is more efficient in correcting character errors than a random system but shows an

average improvement of only about 0.5 db SNR, and

6 - The single null system realizes only a small fraction of the potential improvement of 3 db obtainable from a system which successfully corrects all single bit parity errors.

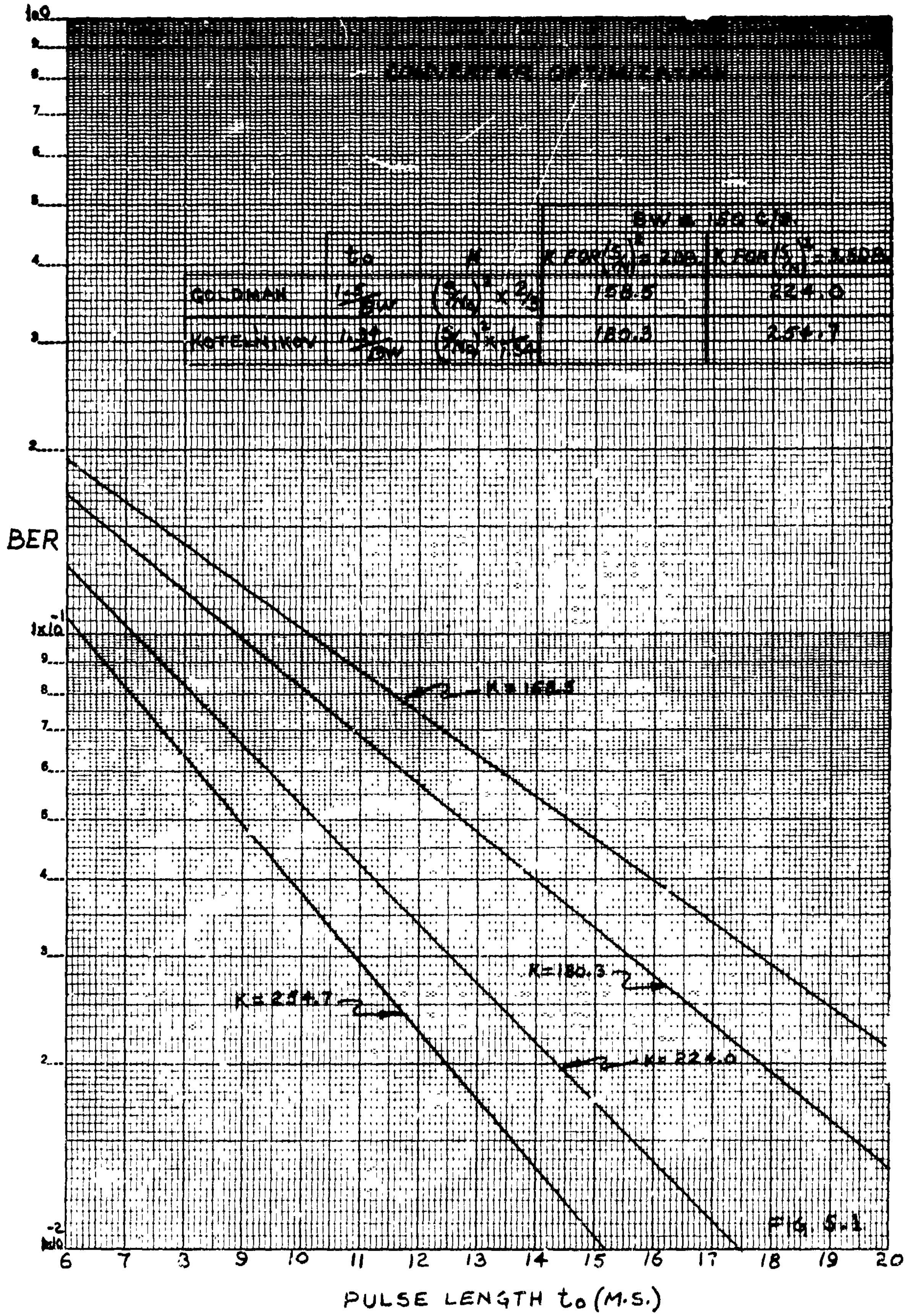

BER
GOLDMAN
KOTELNIKOV
168.5
224.0
180.3
254.7
K=168.5
K=180.3
K=254.7
K=224.0
6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
PULSE LENGTH t_o (M.S.)

FIG. 5-1

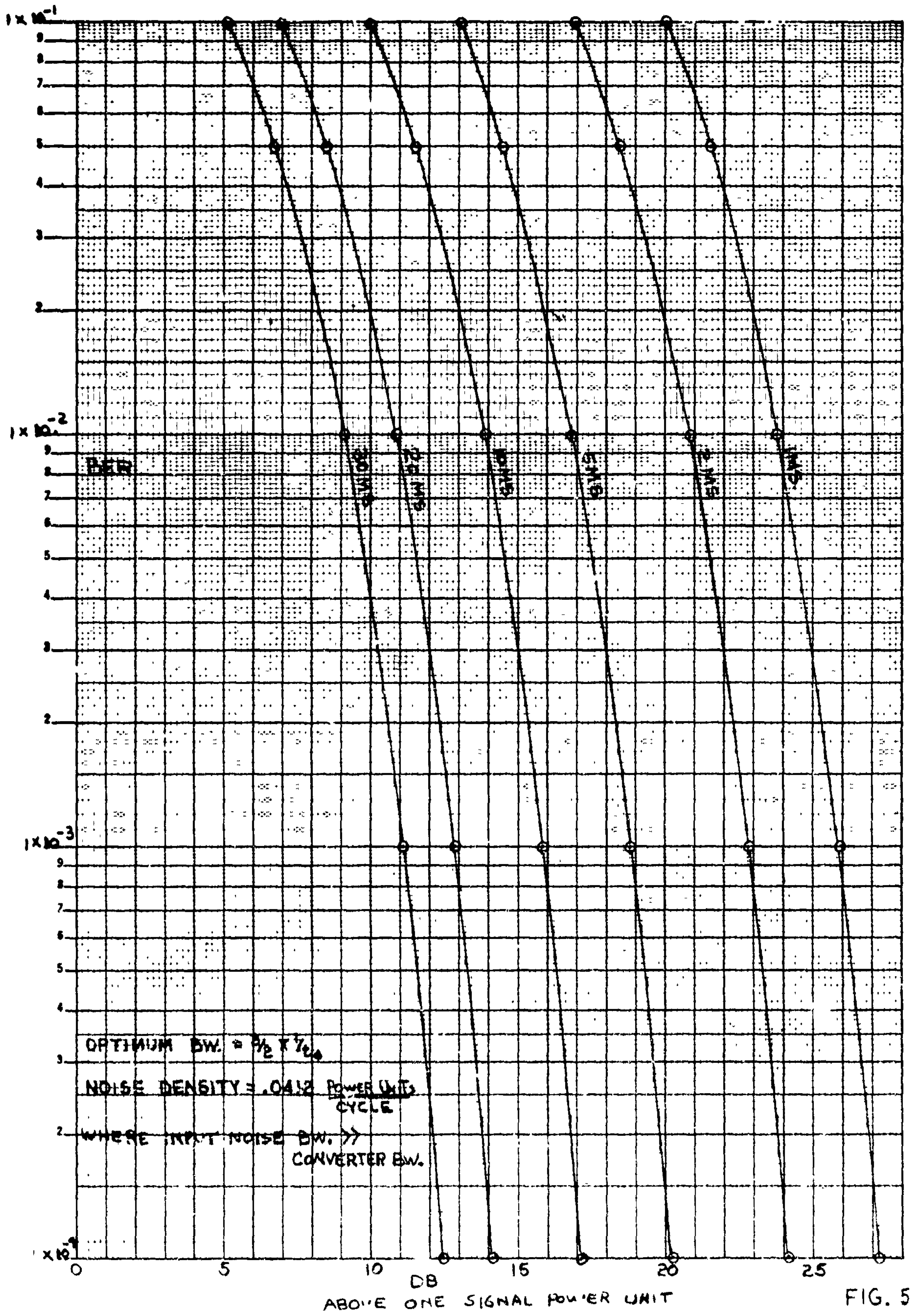

BER
30 MS
20 MS
10 MS
5 MS
2 MS
1 MS
OPTIMUM BW. = 3/2 X 1/t
NOISE DENSITY = .0412 POWER UNITS / CYCLE
WHERE INPUT NOISE BW. >> CONVERTER BW.
0
5
10
15
20
25
DB
ABOVE ONE SIGNAL POWER UNIT

FIG. 5.2

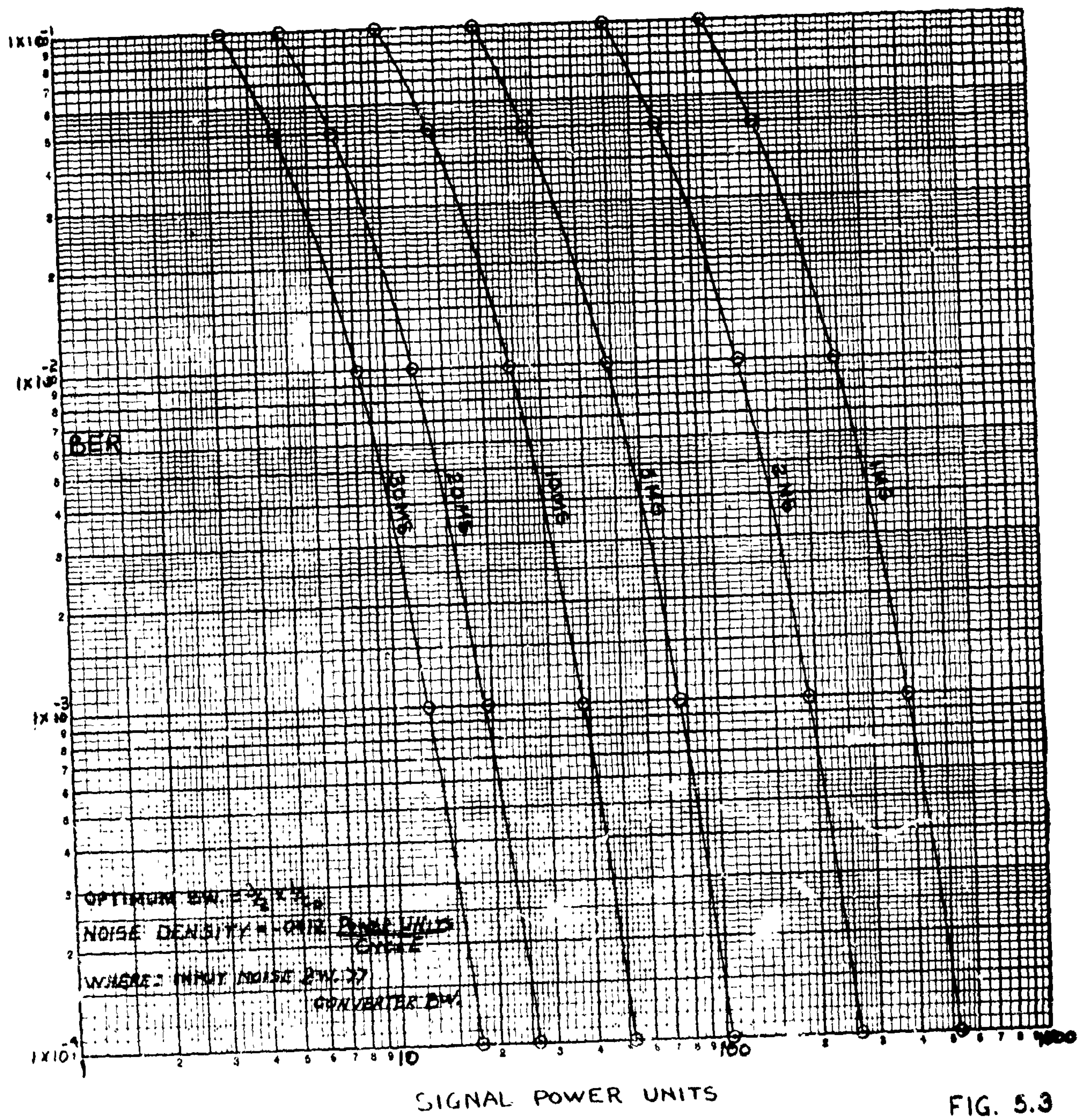
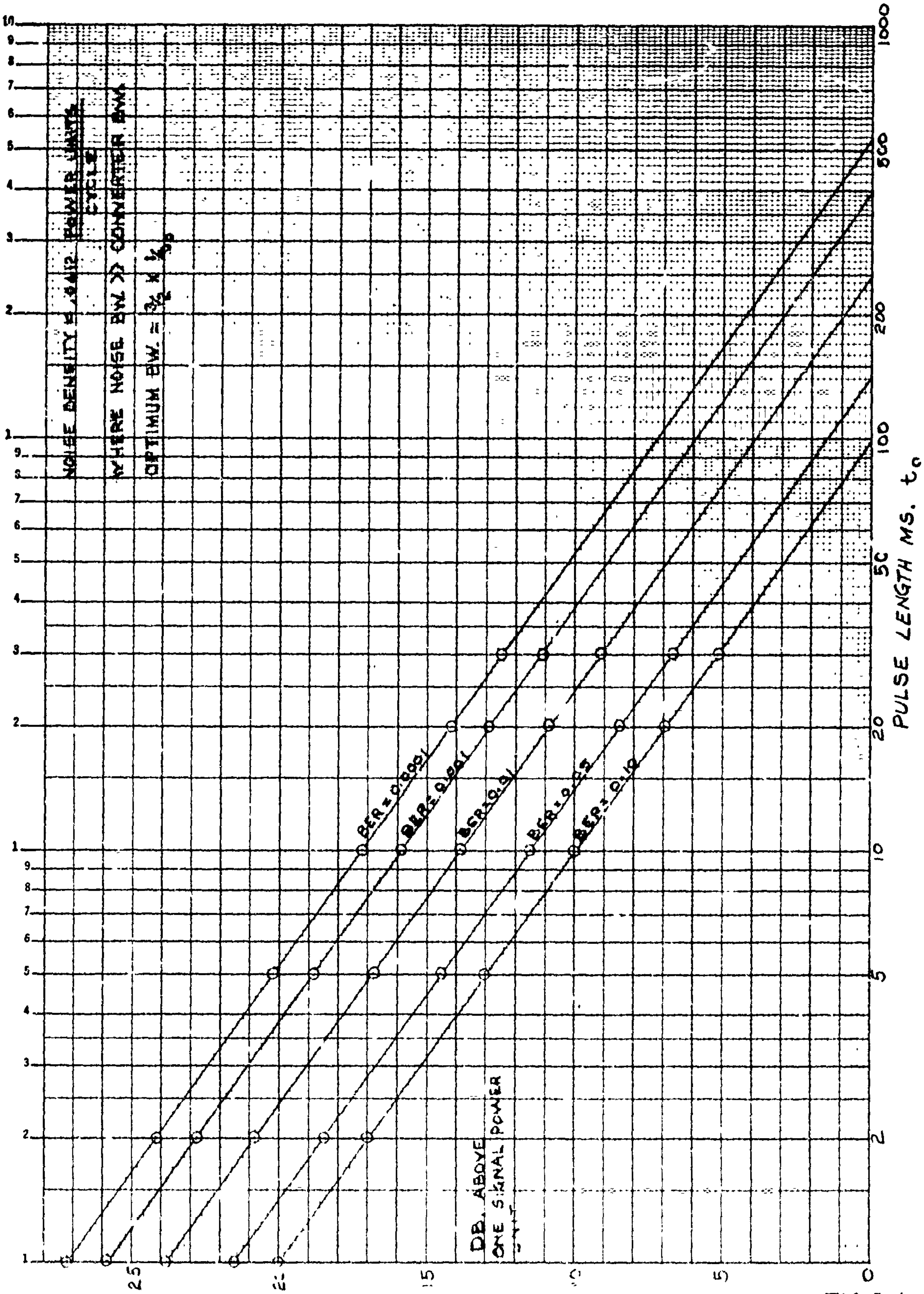

BER
30 MHS
20 MHS
10 MHS
5 MHS
2 MHS
1 MHS
NOISE DENSITY
WHERE: INPUT NOISE BW.
CONVERTER BW.
SIGNAL POWER UNITS

FIG. 5.3



NOISE DENSITY = .0412 POWER UNITS / CYCLE
WHERE NOISE BW >> CONVERTER BW
OPTIMUM BW = 3/2 x 1/t0
BER = 0.0001
BER = 0.001
BER = 0.01
BER = 0.05
BER = 0.10
DB. ABOVE ONE SIGNAL POWER
PULSE LENGTH MS. t0

FIG 5.4

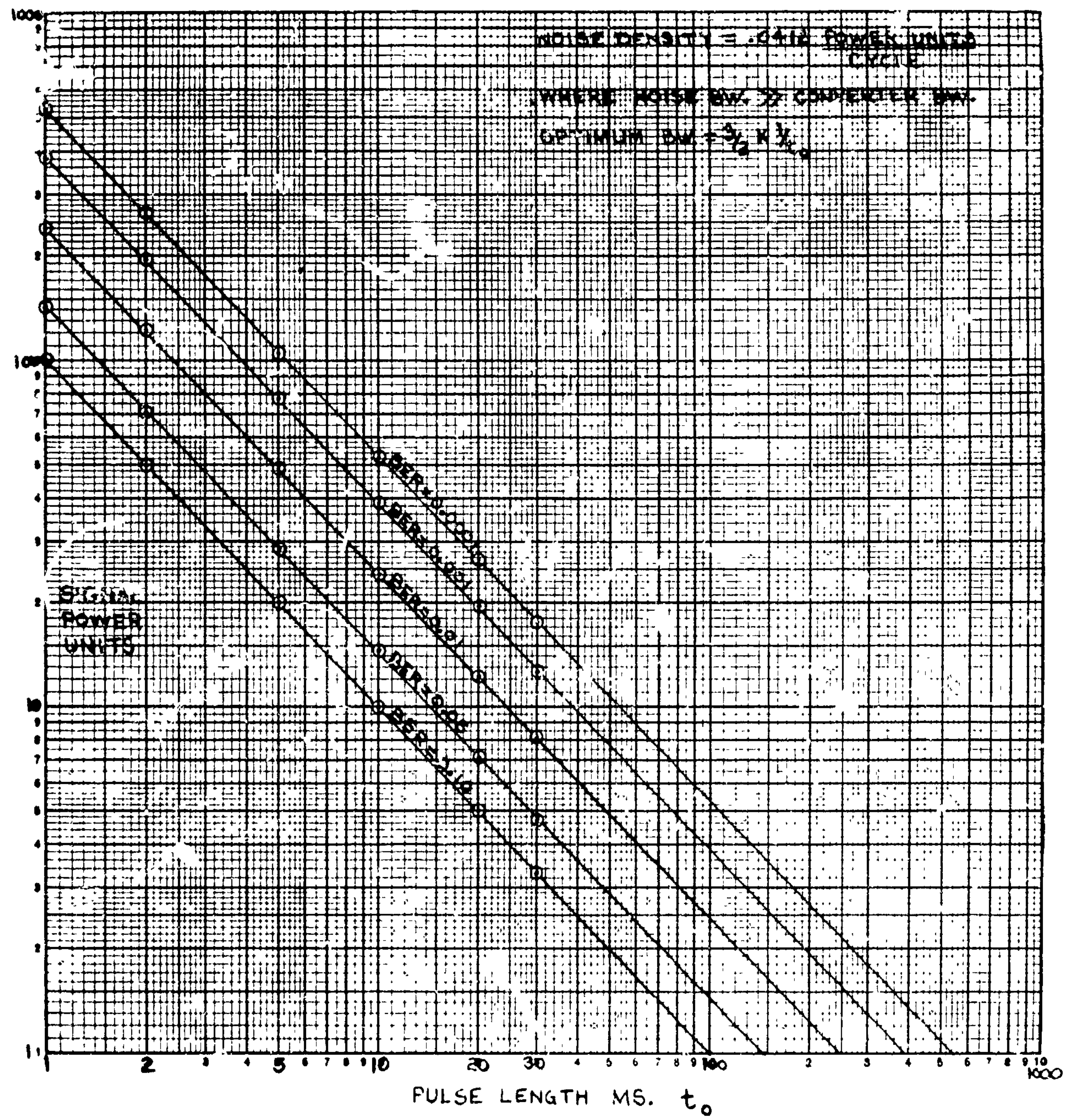

NOISE DENSITY = .0412 POWER UNITS / CYCLE
WHERE NOISE BW. >> CONVERTER BW.
OPTIMUM BW = $\frac{3}{2} K \frac{1}{t_o}$
SIGNAL POWER UNITS
PULSE LENGTH MS. t_o

FIG. 5.5

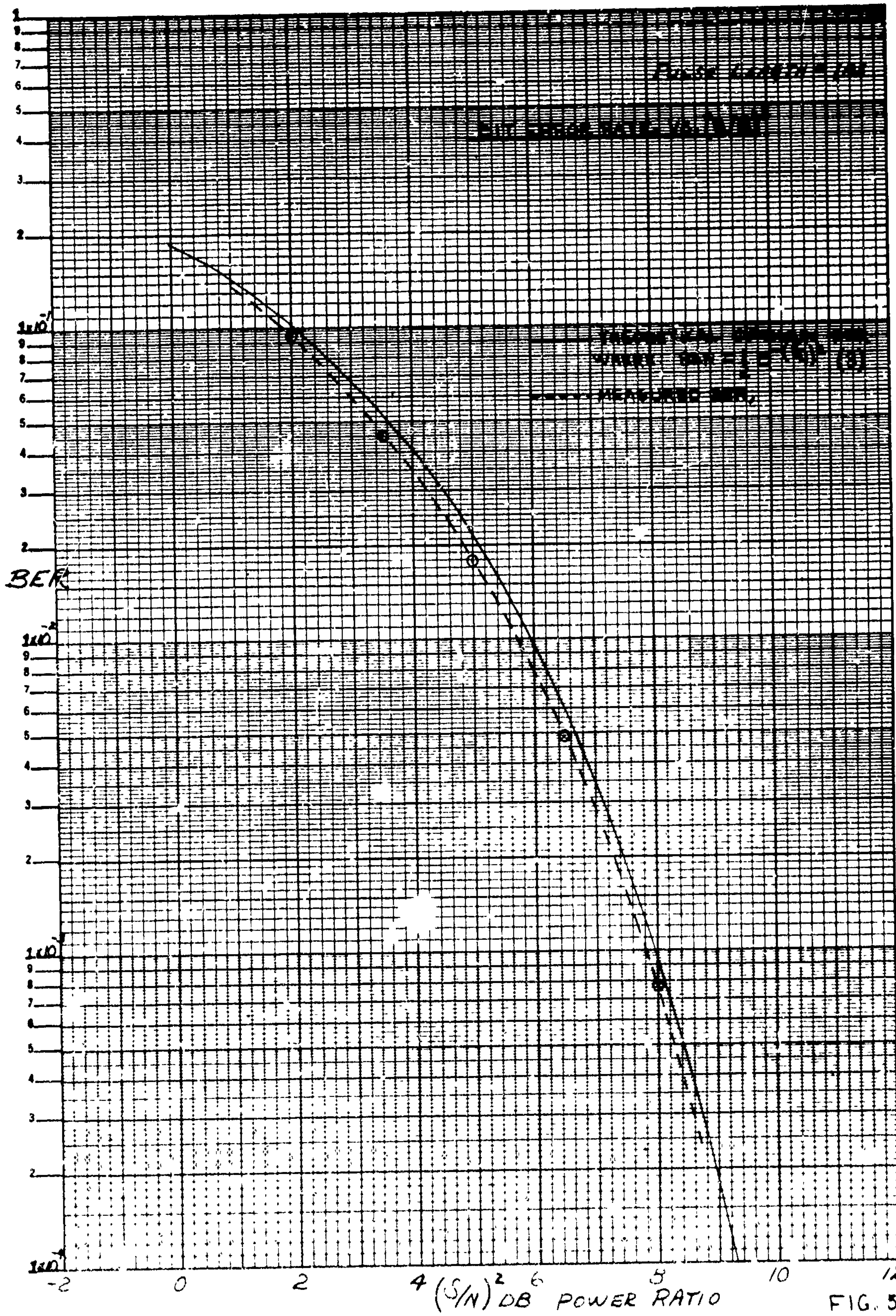

BER
1
9
8
7
6
5
4
3
2
1x10⁻¹
1x10⁻²
1x10⁻³
1x10⁻⁴
-2
0
2
4
6
8
10
12
(S/N)² DB POWER RATIO

FIG. 5.6A

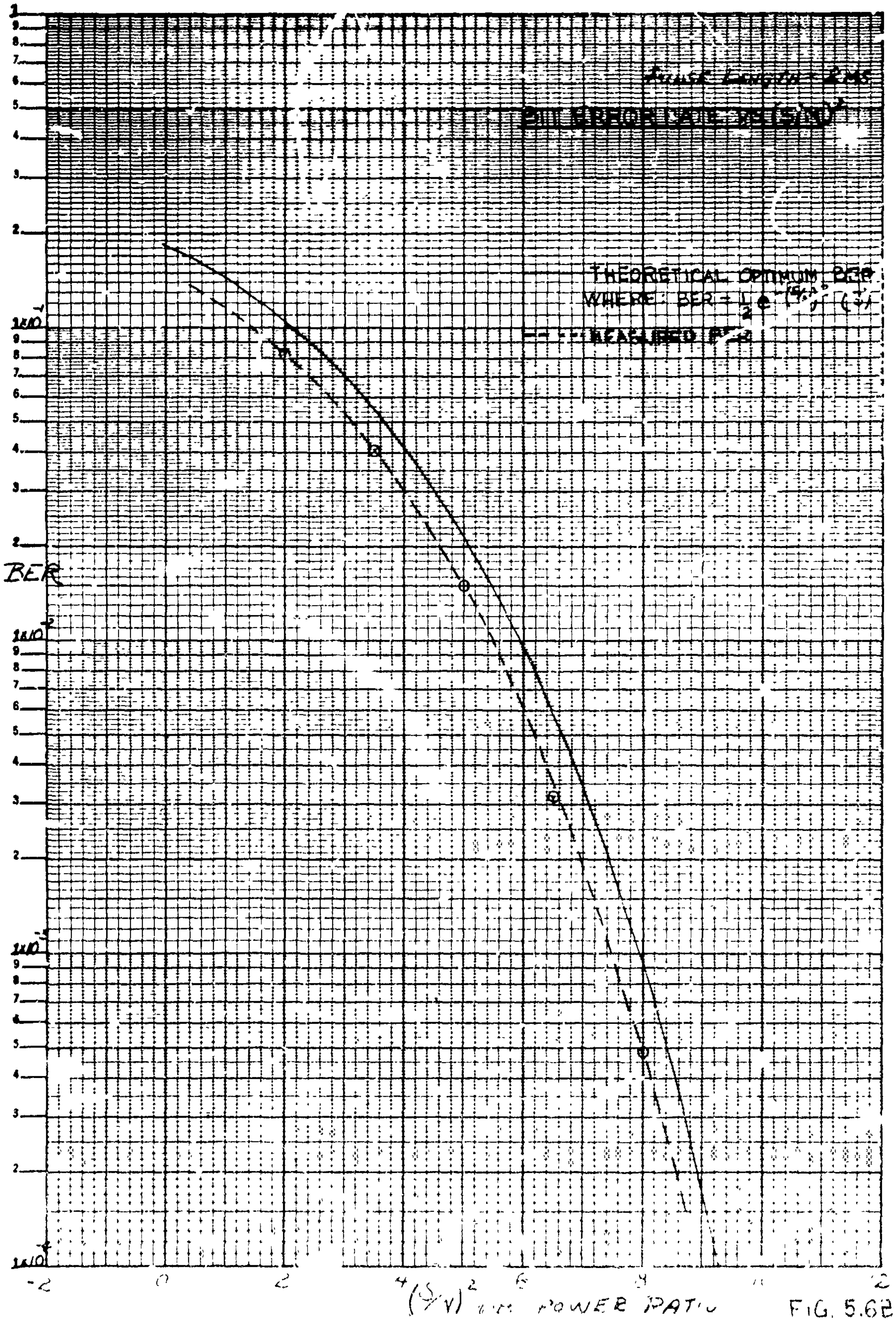


FIG. 5.62

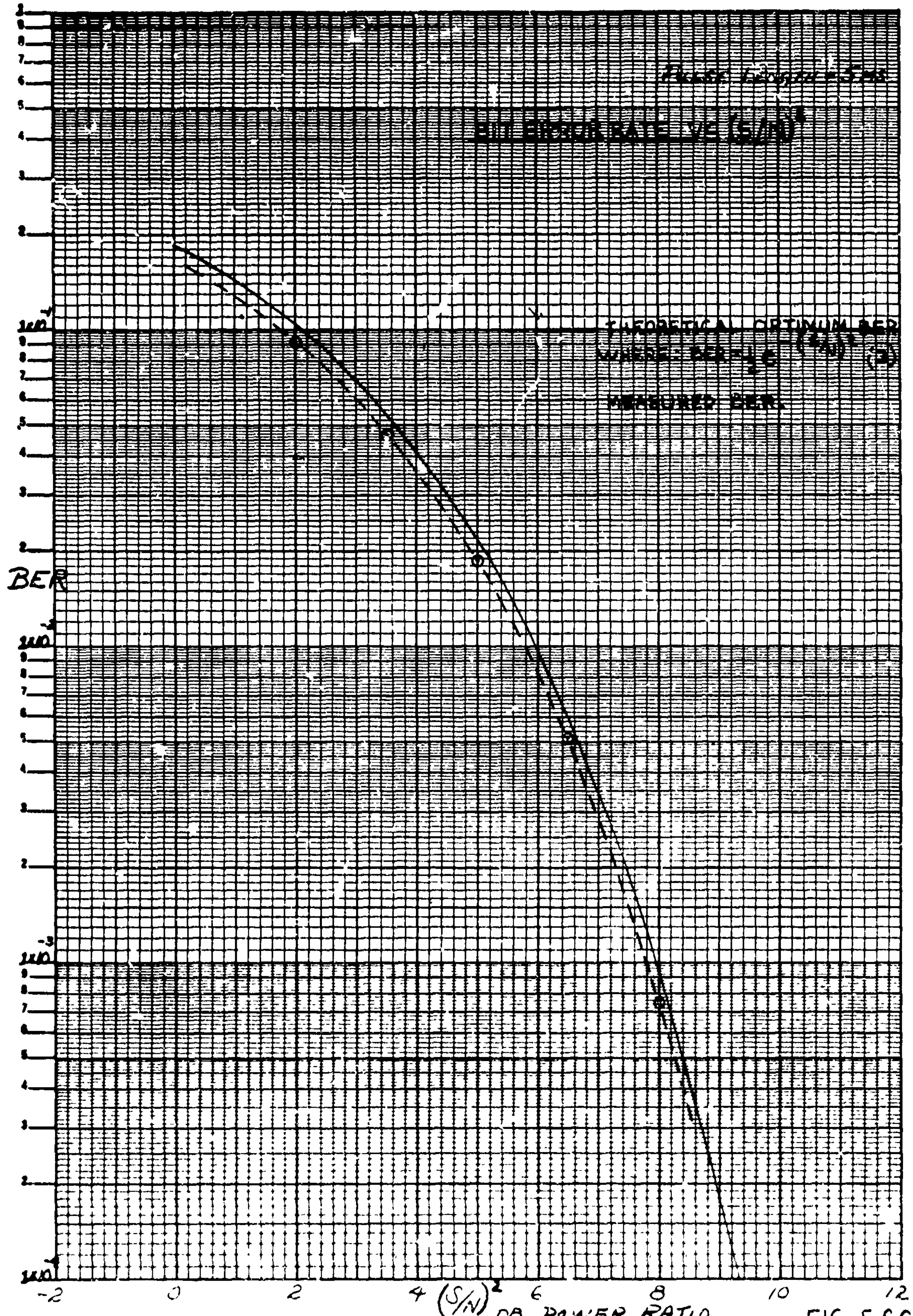

BIT ERROR RATE VS (S/N)²
THEORETICAL OPTIMUM BER
MEASURED BER
BER
-2
0
2
4
6
8
10
12
(S/N)² DB POWER RATIO

FIG. 5.6C

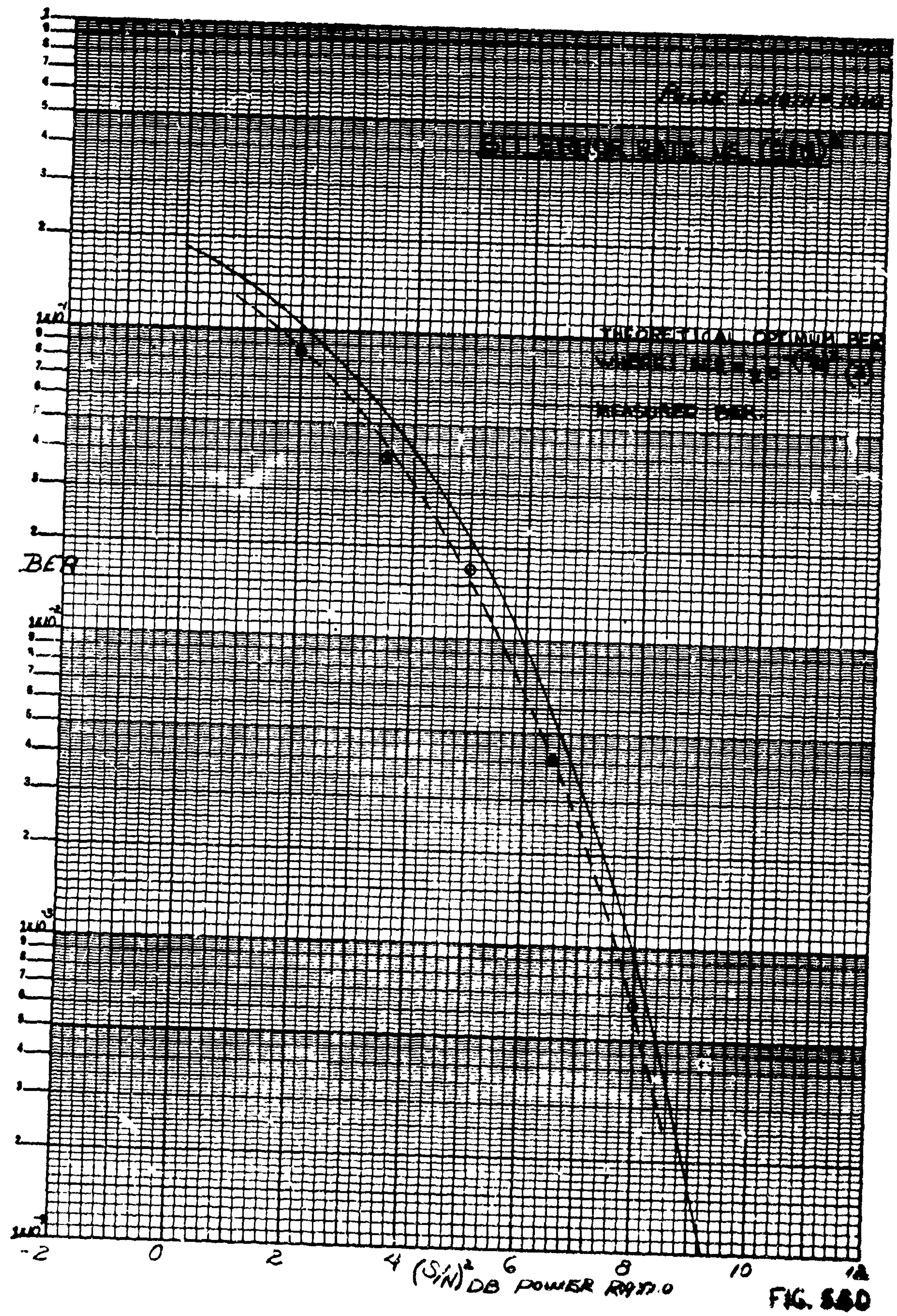

THEORETICAL OPTIMUM BER
MEASURED BER
BER
-2
0
2
4
6
8
10
12
$(S/N)^2$ DB POWER RATIO

FIG. 5.8D

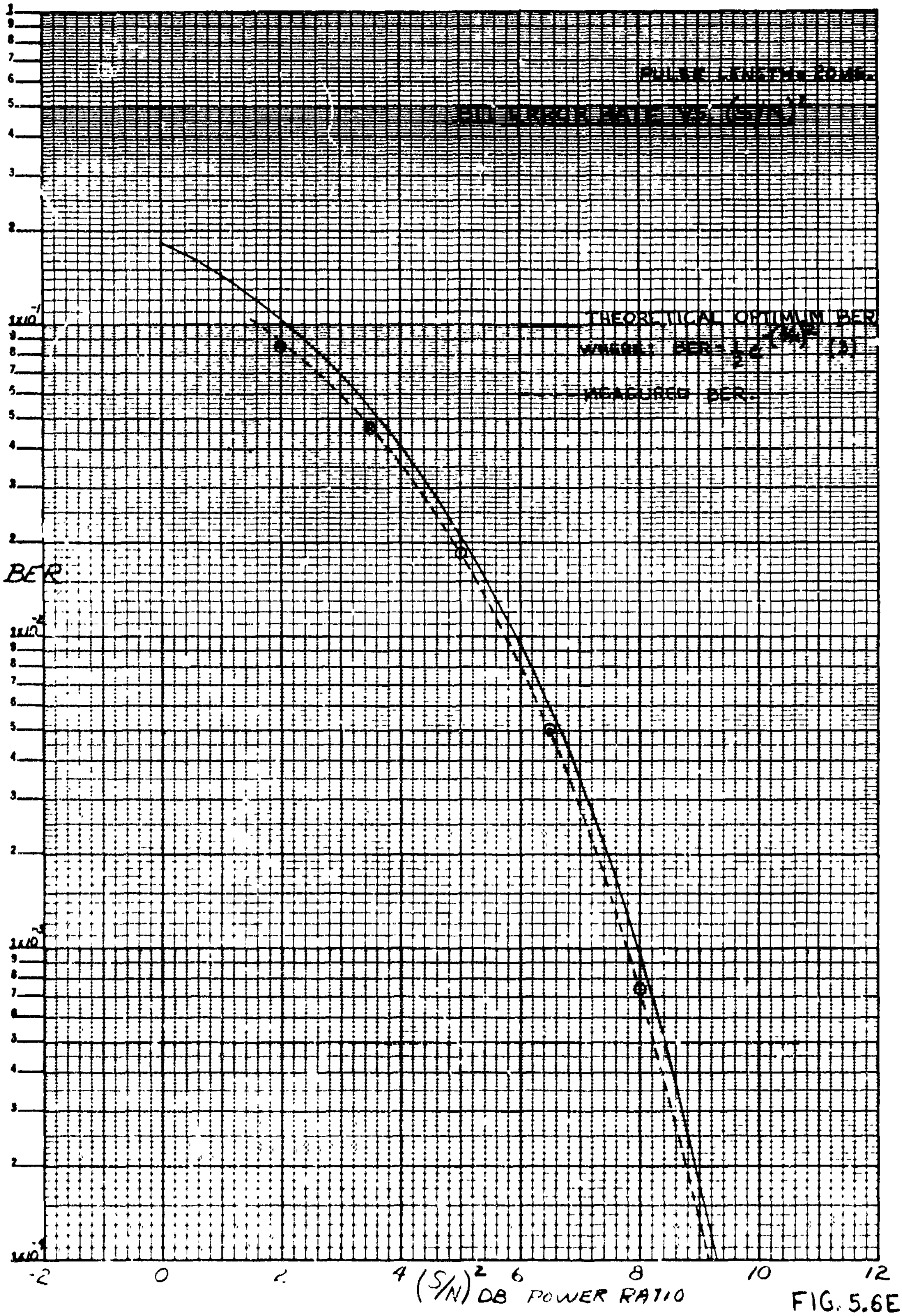

THEORETICAL OPTIMUM BER
MEASURED BER.
BER
-2
0
2
4
6
8
10
12
(S/N)² DB POWER RATIO

FIG. 5.6E

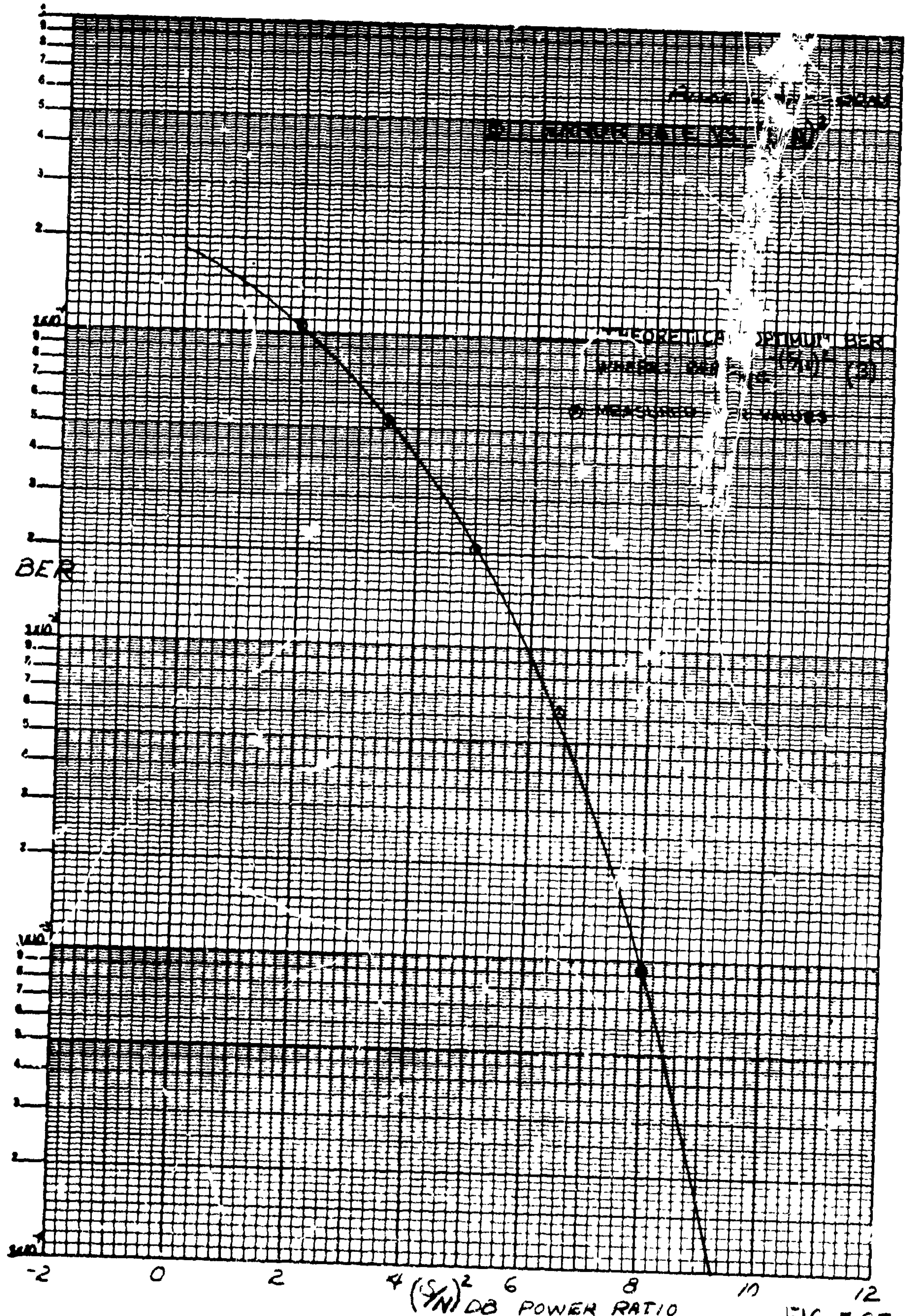


FIG. 5.8F

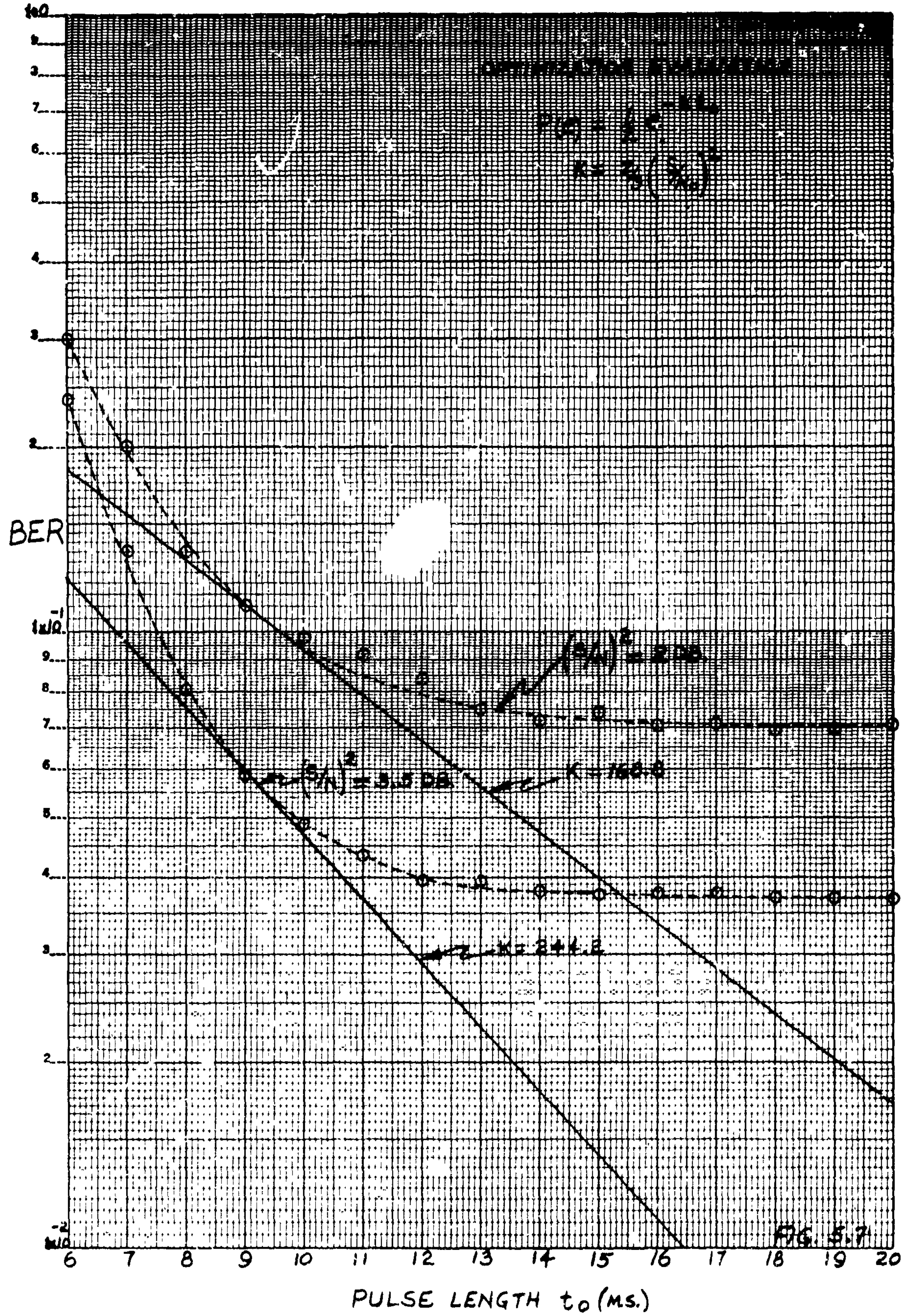
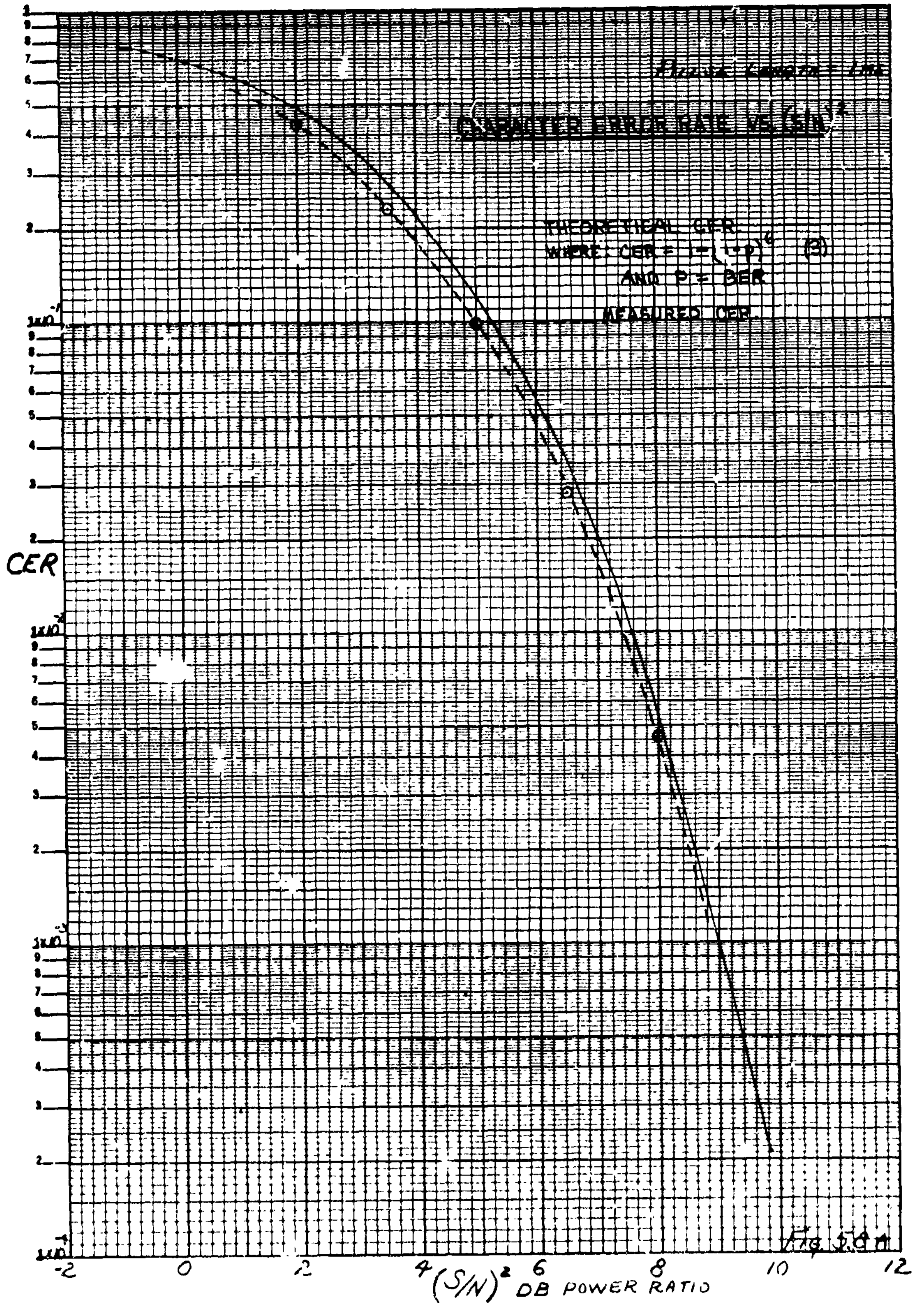
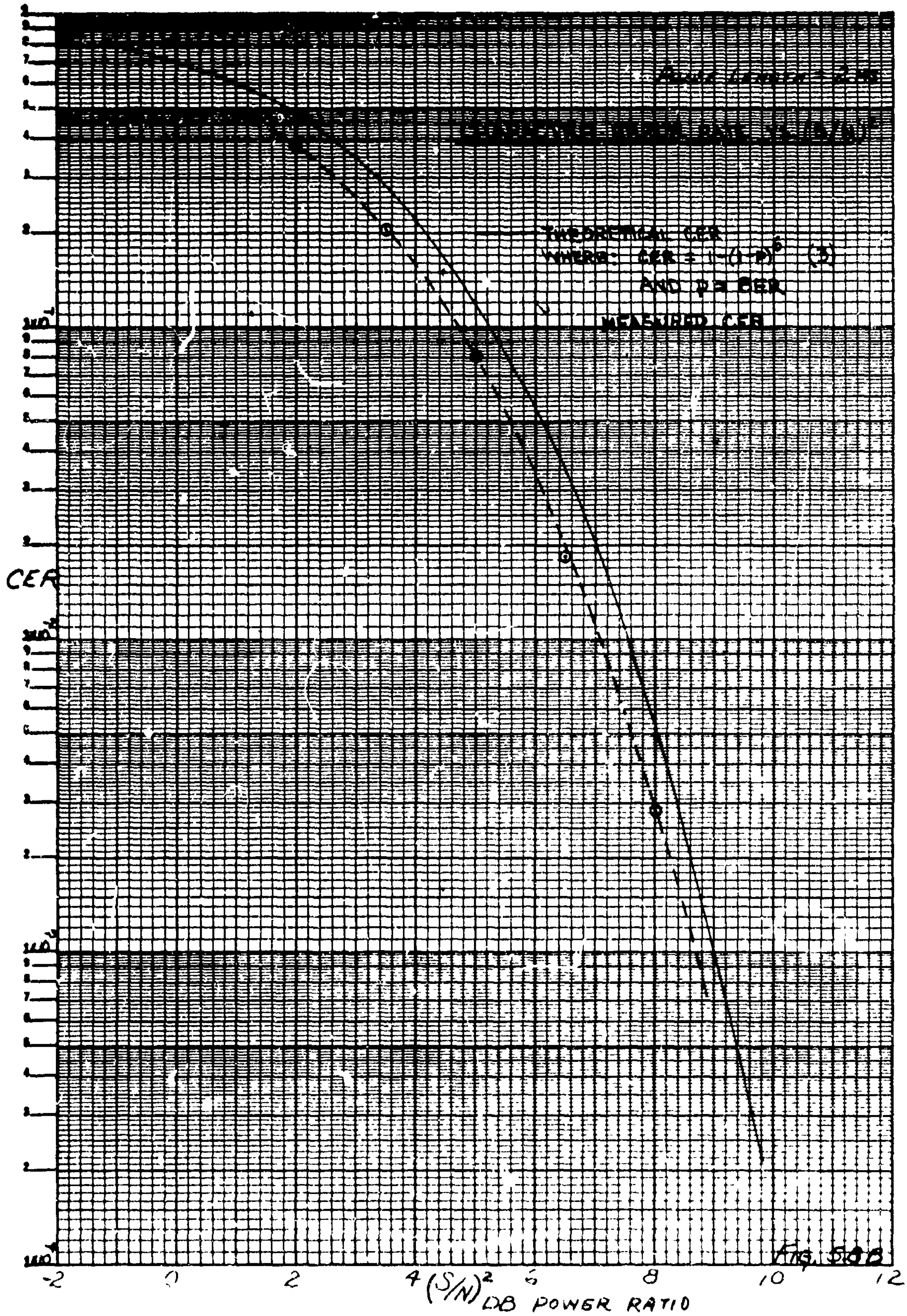

BER
PULSE LENGTH to (MS.)
K = 160.8
K = 241.2
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20

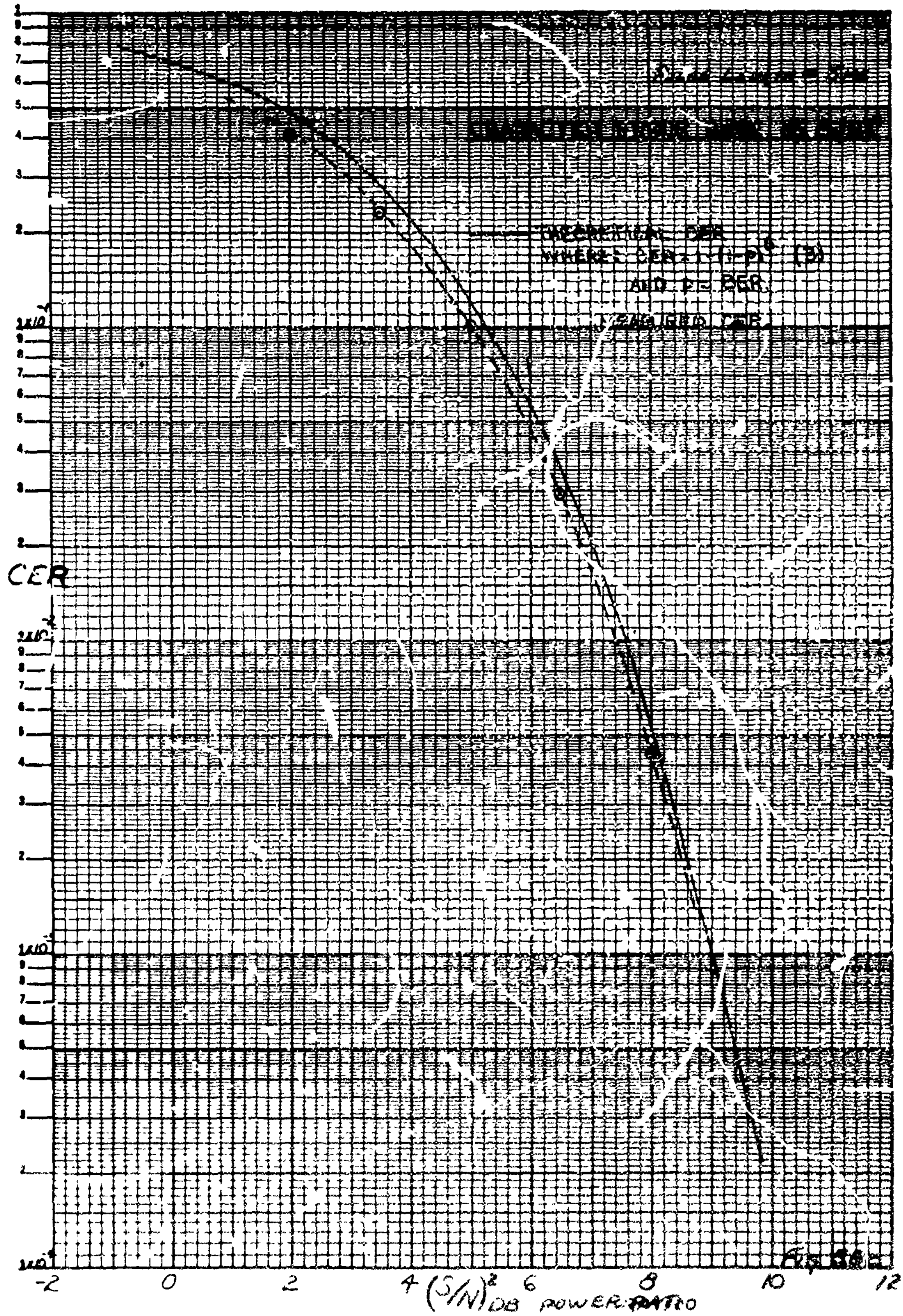
FIG. 5.7

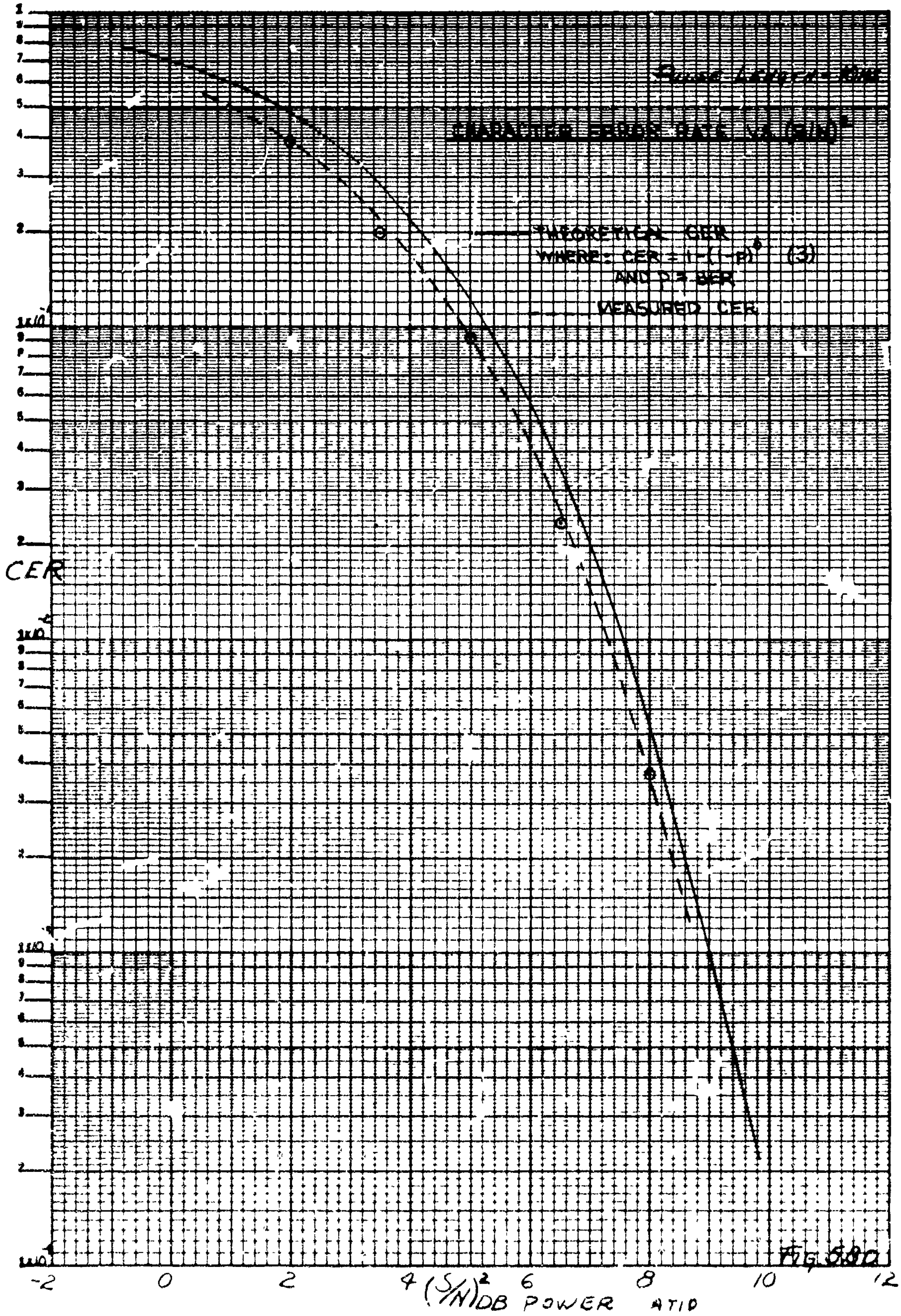


CHARACTER ERROR RATE VS (S/N)^2
THEORETICAL CER
WHERE: CER = 1-(1-P)^6 (3)
AND P = BER
MEASURED CER
CER
(S/N)^2 DB POWER RATIO



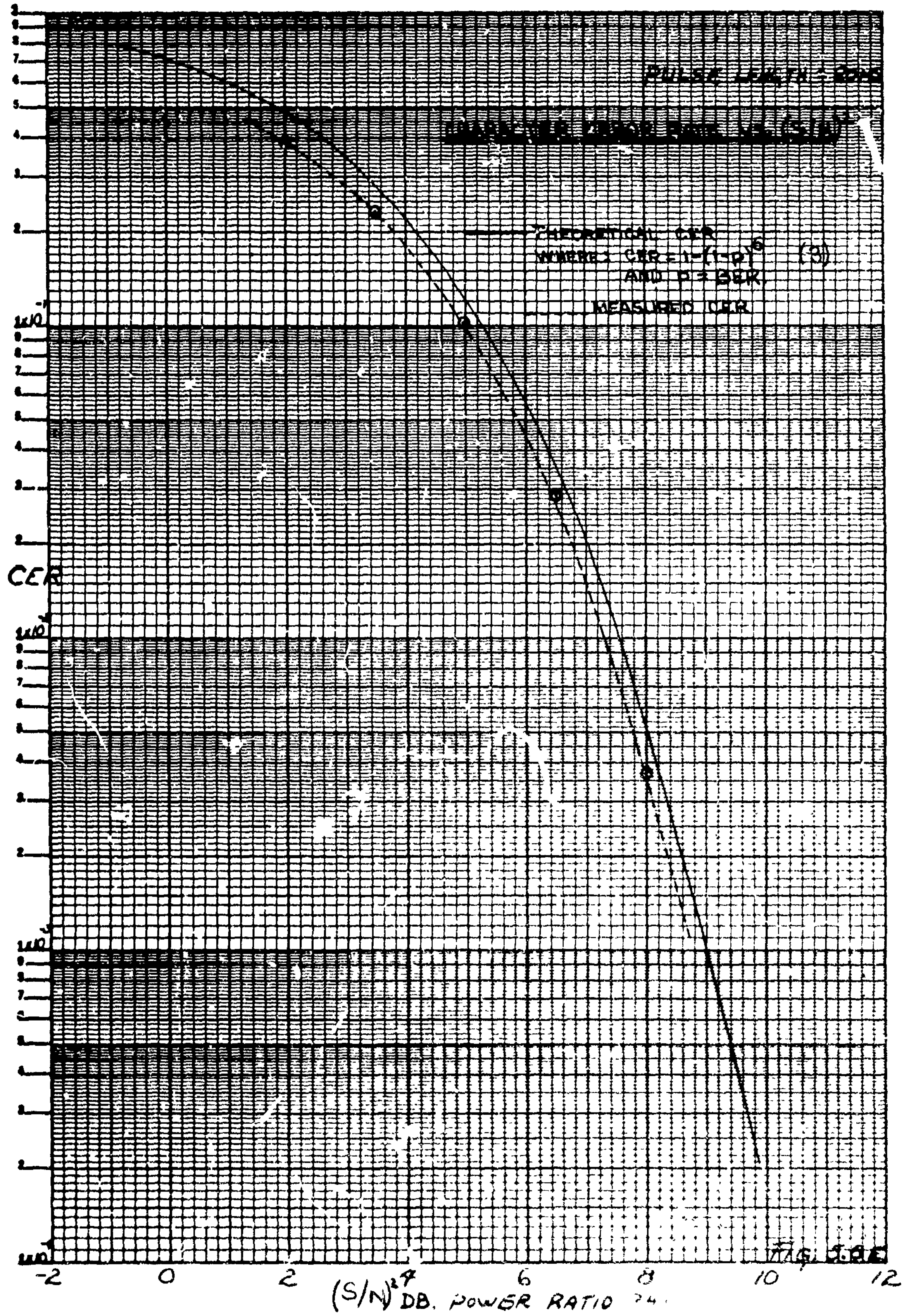
THEORETICAL CER
WHERE: CER = 1-(1-P)^6 (3)
AND P = BER
MEASURED CER
CER
-2
0
2
4
6
8
10
12
(S/N)² DB POWER RATIO
Fig. 50B

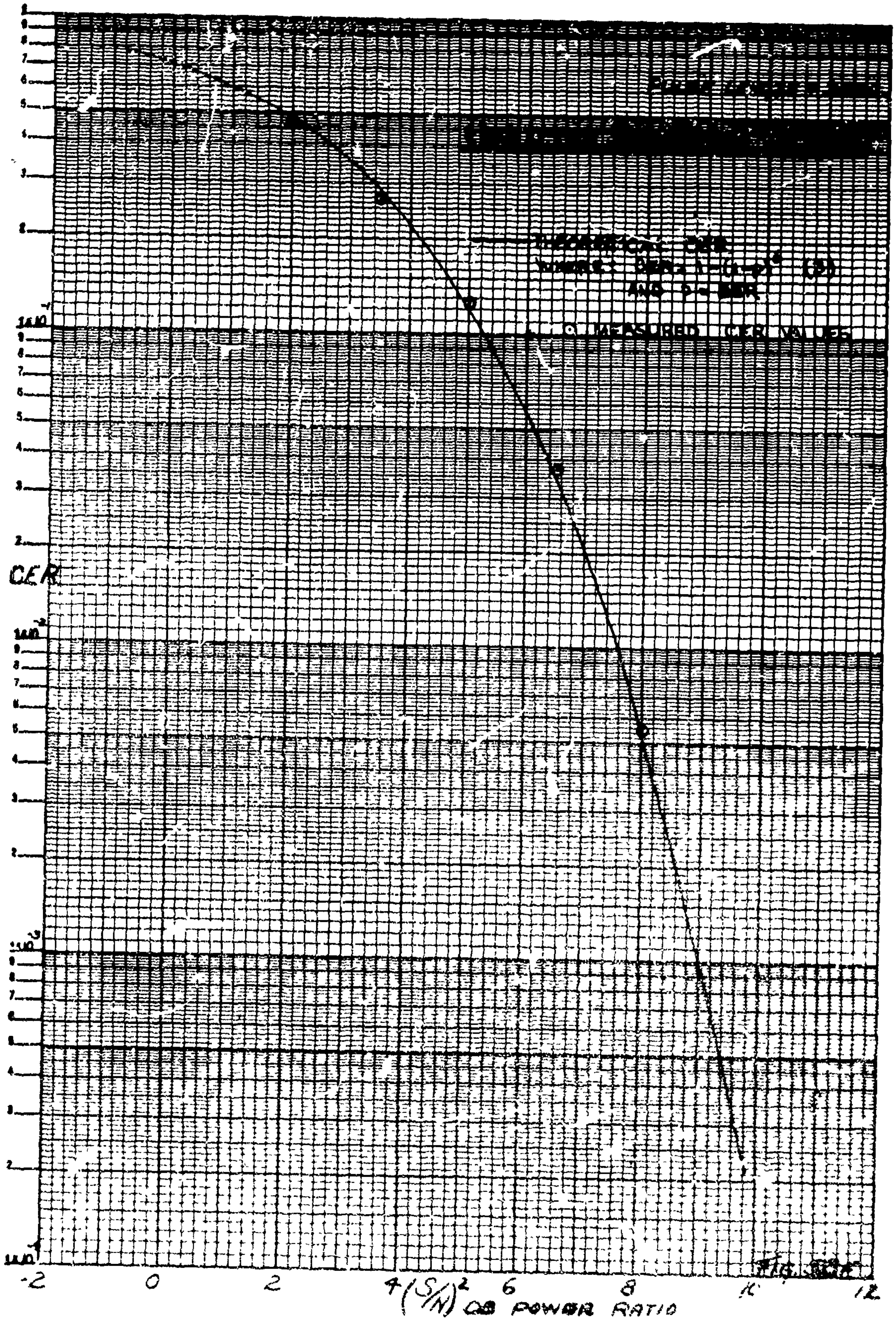

CER
AND P = BER
-2
0
2
4
6
8
10
12
$(S/N)^2_{DB}$ POWER RATIO



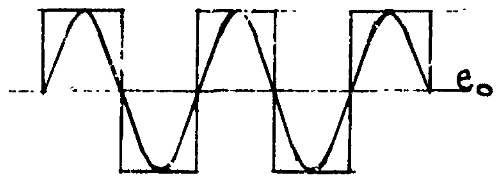
CHARACTER ERROR RATE VS (S/N)²
THEORETICAL CER
WHERE: CER = 1-(1-P)⁶ (3)
AND P = BER
MEASURED CER
CER
(S/N)² DB POWER RATIO
-2
0
2
4
6
8
10
12

Fig. 58D


THEORETICAL CER
WHERE: CER = 1-(1-P)^6 (3)
AND P = BER
MEASURED CER
CER
1×10^-1
1×10^-2
1×10^-3
1×10^-4
-2
0
2
4
6
8
10
12
(S/N)² DB. POWER RATIO



O MEASURED CER VALUES
CER
-2
0
2
4
6
8
10
12
(S/N)² DB POWER RATIO



DECISION AND BINARY DRIVING WAVEFORMS.

FIG. 5.9 A

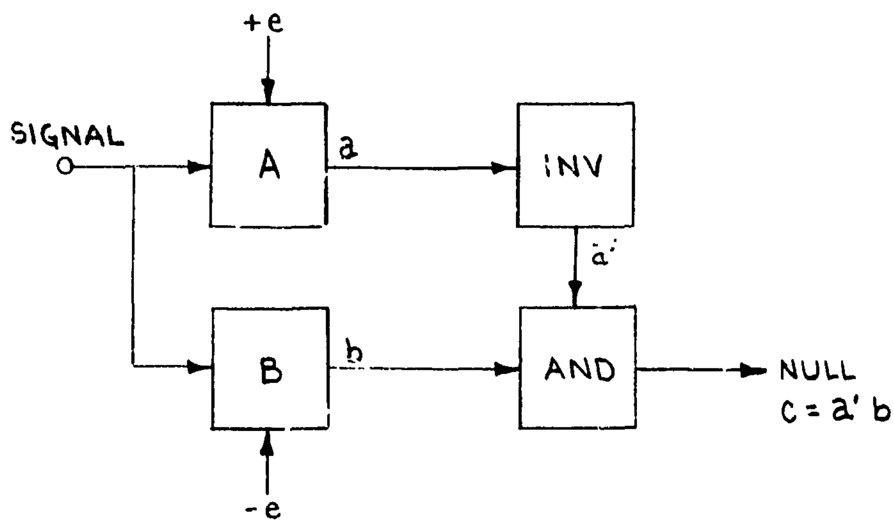


FIG. 5.9 B

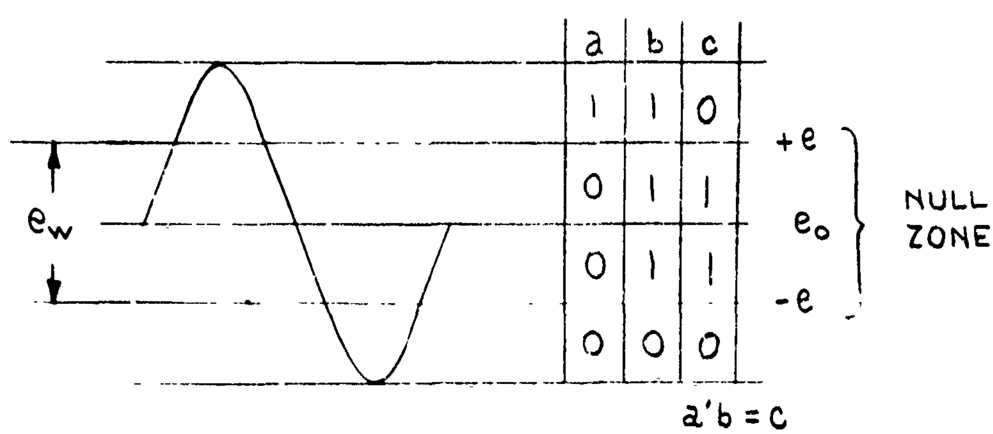


FIG. 5.9 C

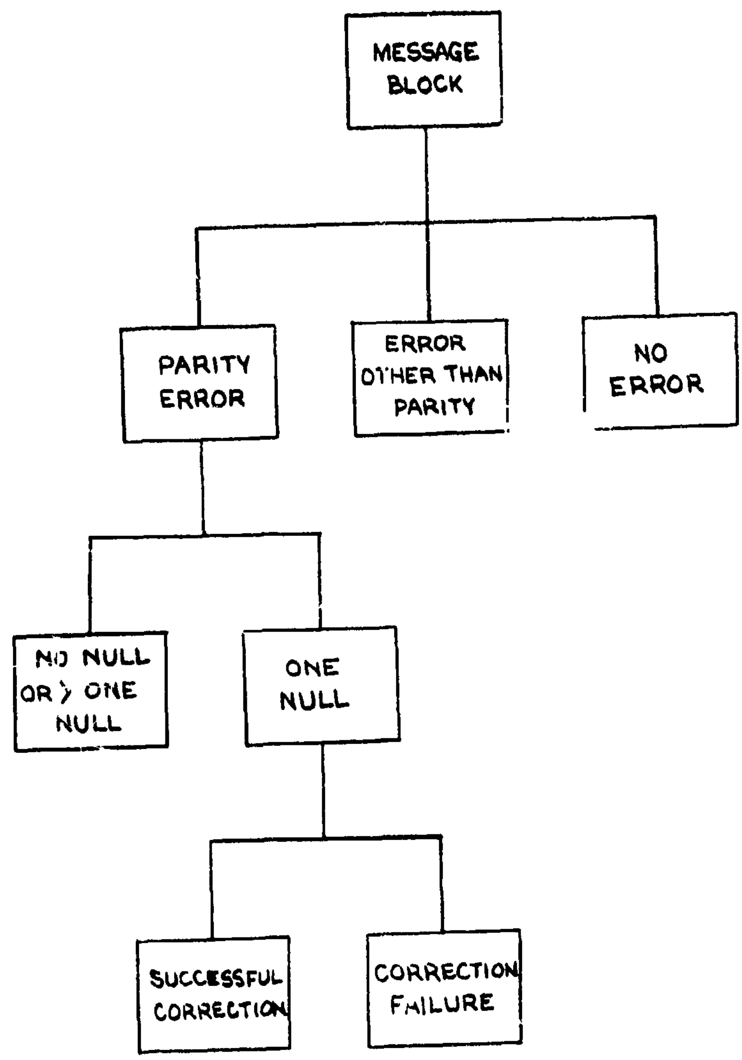
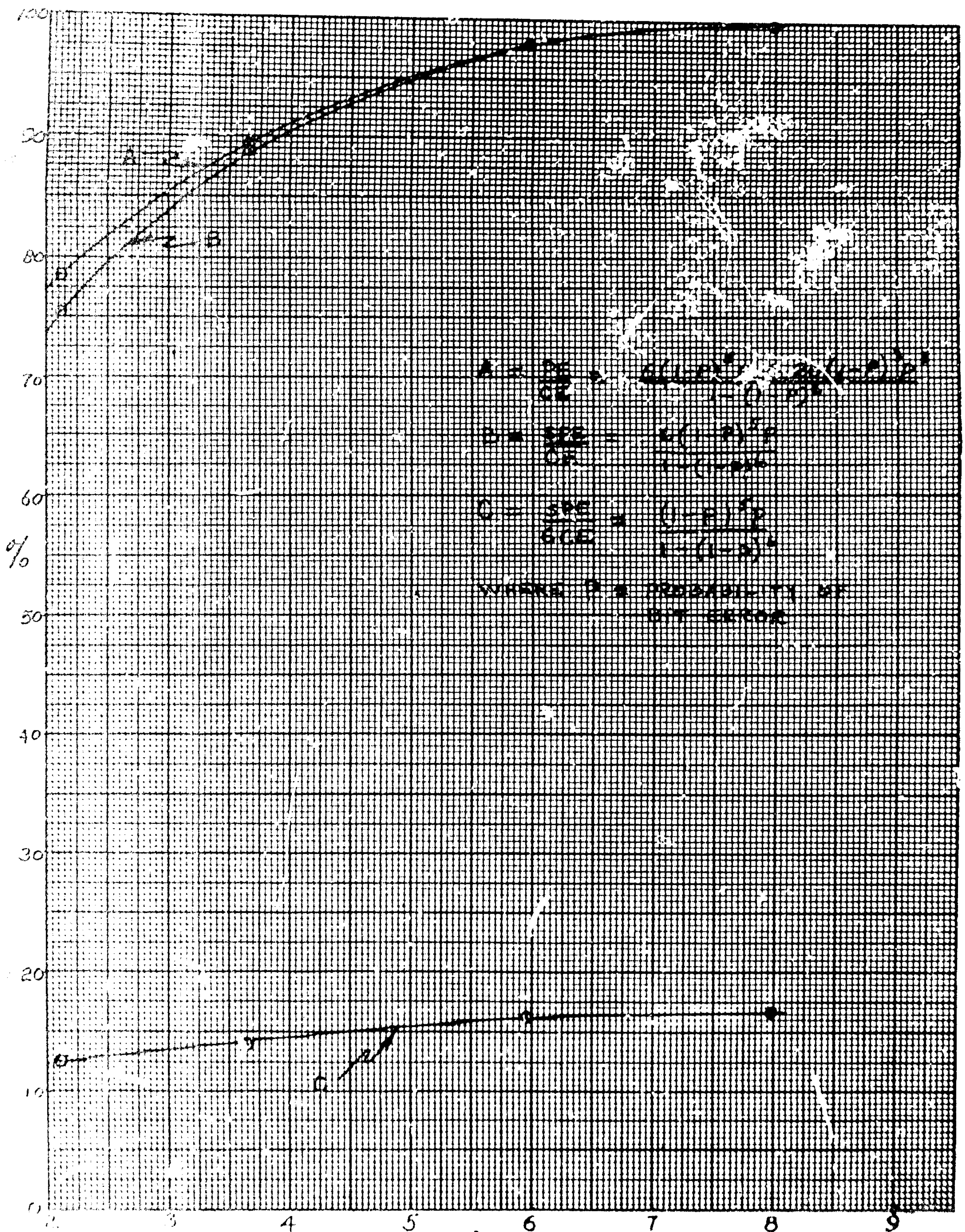

MESSAGE BLOCK
PARITY ERROR
ERROR OTHER THAN PARITY
NO ERROR
NO NULL OR > ONE NULL
ONE NULL
SUCCESSFUL CORRECTION
CORRECTION FAILURE

FIG. 5.9 D



$(S/N)^2$ DB. POWER RATIO

FIG. 5.10A

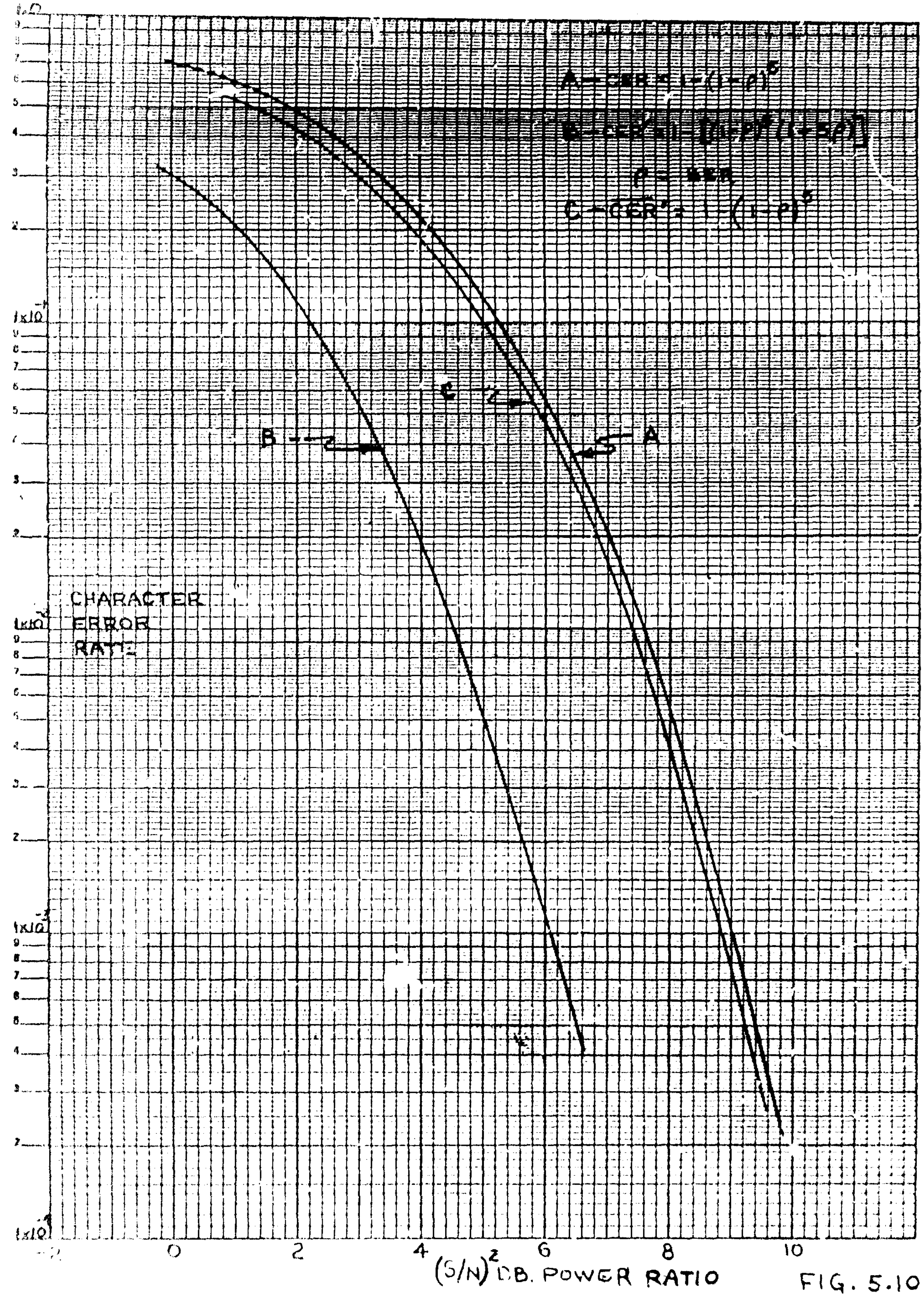
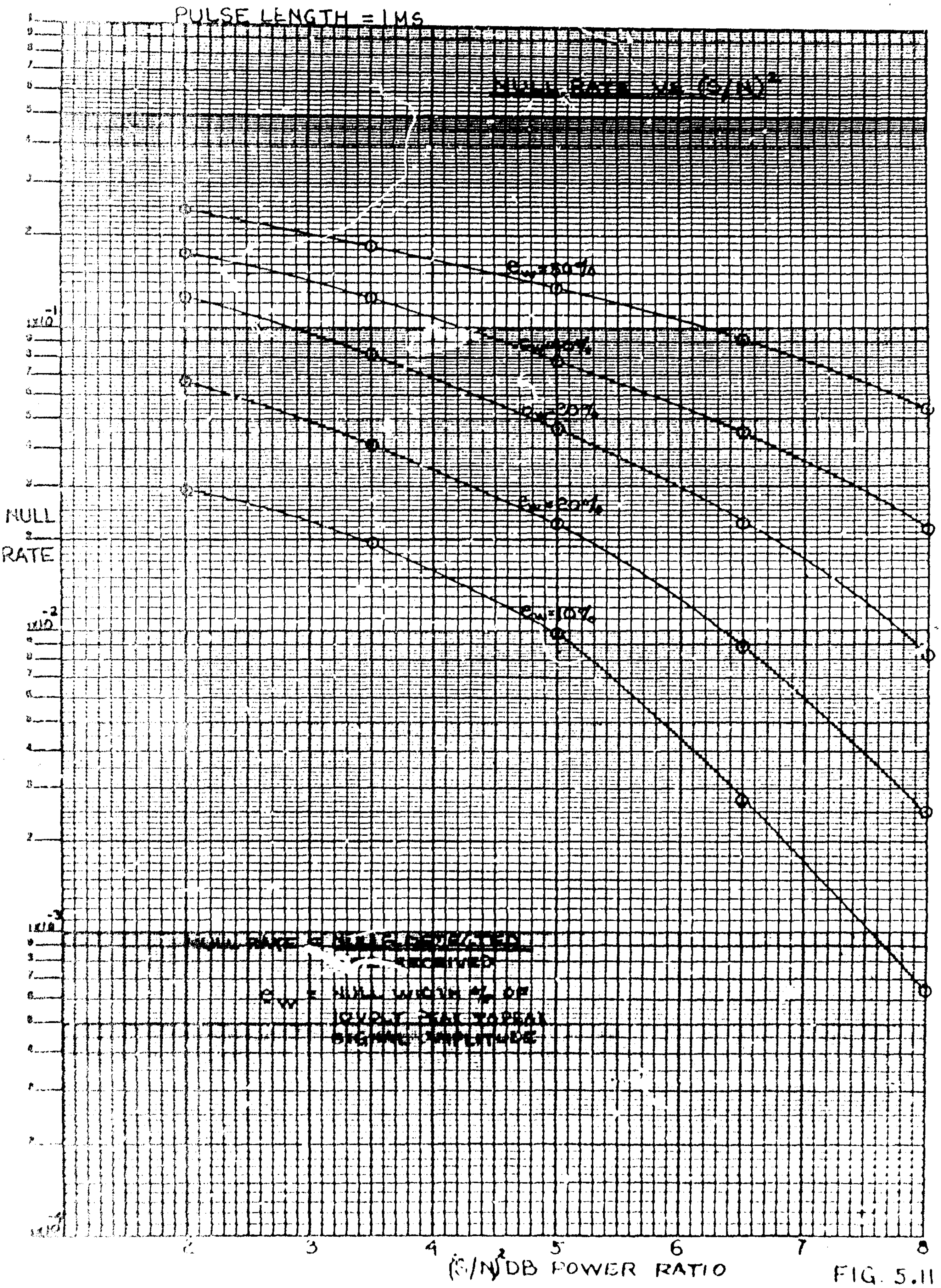


FIG. 5.10 B



PULSE LENGTH = 1MS
NULL RATE VS (S/N)²
NULL RATE
1x10⁻¹
1x10⁻²
1x10⁻³
eW=20%
eW=10%
2
3
4
5
6
7
8
(S/N)² DB POWER RATIO

FIG. 5.11 A

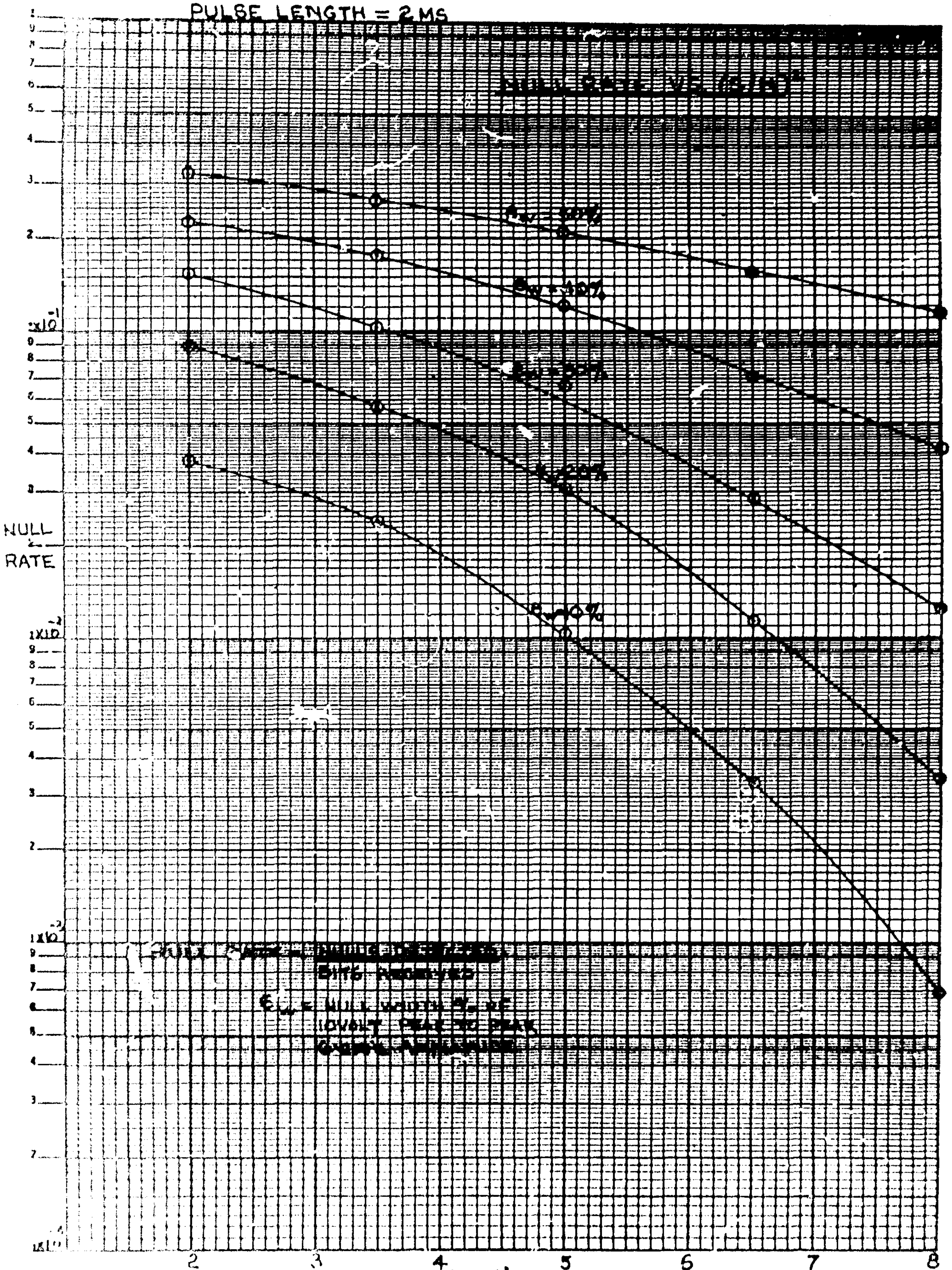

PULSE LENGTH = 2 MS
NULL RATE
1×10^{-1}
1×10^{-2}
1×10^{-3}
1×10^{-4}
$\epsilon_w = 50\%$
$\epsilon_w = 40\%$
$\epsilon_w = 30\%$
$\epsilon_w = 20\%$
$\epsilon_w = 10\%$
BITS RECEIVED
10 VOLT PEAK TO PEAK
2
3
4
5
6
7
8
$(S/N)^2$ DB POWER RATIO

FIG. 5.11B

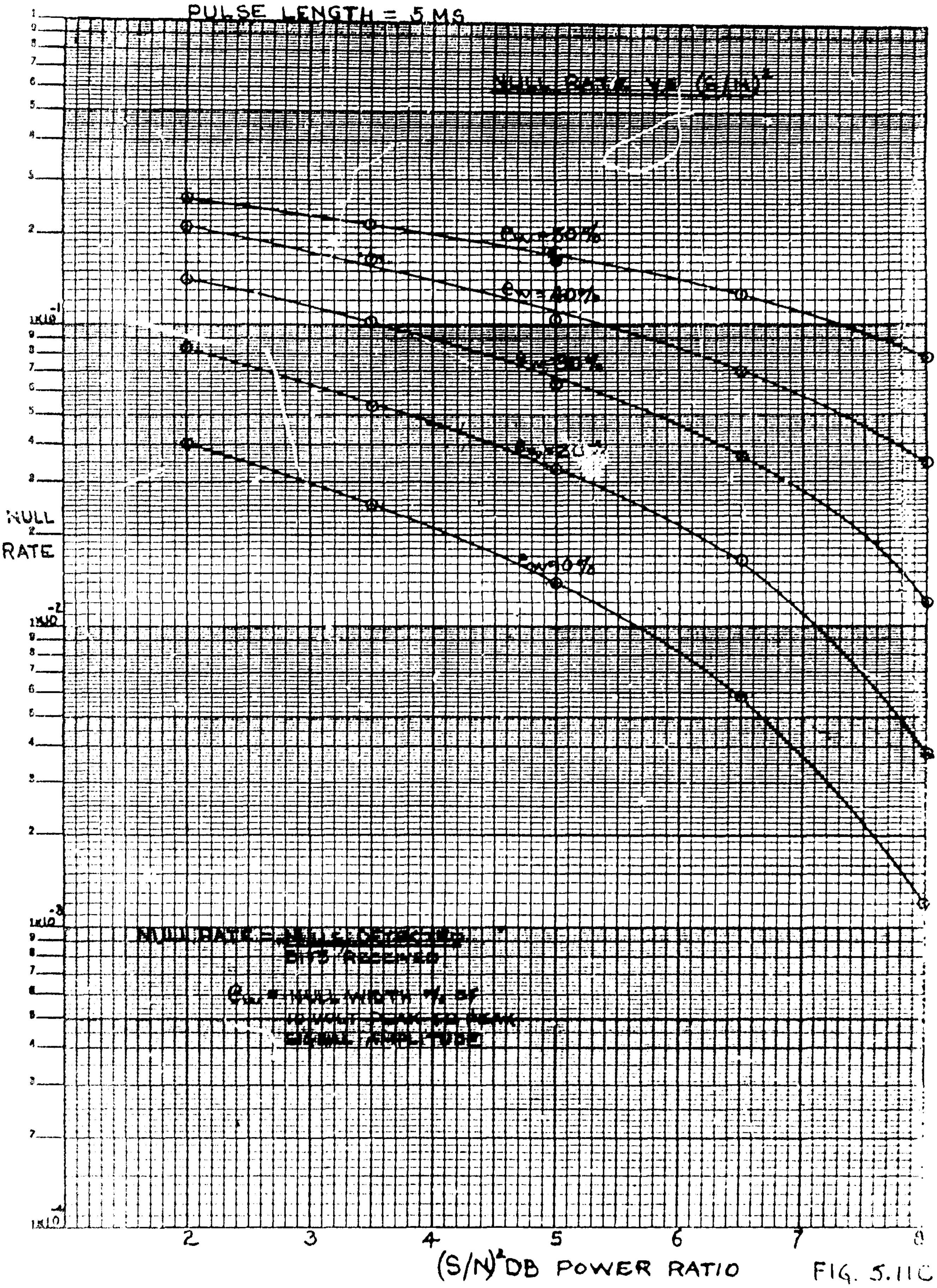

PULSE LENGTH = 5 MS
NULL RATE VS (S/N)²
e_w = 50%
e_w = 40%
e_w = 30%
e_w = 20%
e_w = 10%
NULL RATE = NULLS DETECTED / BITS RECEIVED
e_w = NULL WIDTH % OF 10 VOLT PEAK-TO-PEAK SIGNAL AMPLITUDE
NULL RATE
1x10^{-1}
1x10^{-2}
1x10^{-3}
1x10^{-4}
2
3
4
5
6
7
8
(S/N)² DB POWER RATIO

FIG. 5.11C

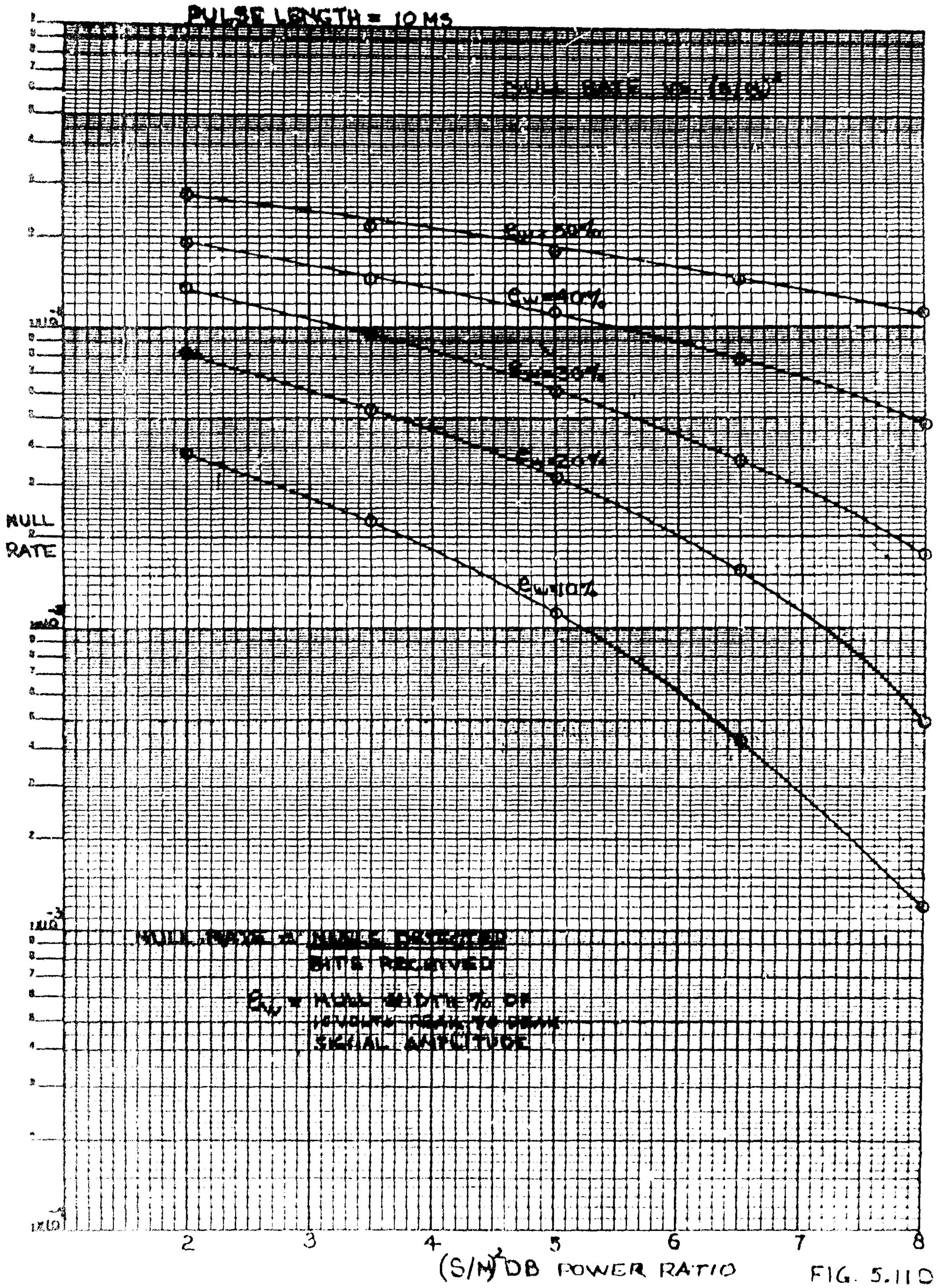

PULSE LENGTH = 10 MS
NULL RATE vs (S/N)²
e_w = 50%
e_w = 40%
e_w = 30%
e_w = 20%
e_w = 10%
NULL RATE
1×10⁻¹
1×10⁻²
1×10⁻³
1×10⁻⁴
NULL RATE = NULLS DETECTED / BITS RECEIVED
e_w = NULL WIDTH % OF
PEAK TO PEAK
SIGNAL AMPLITUDE
2
3
4
5
6
7
8
(S/N)² DB POWER RATIO

FIG. 5.11D

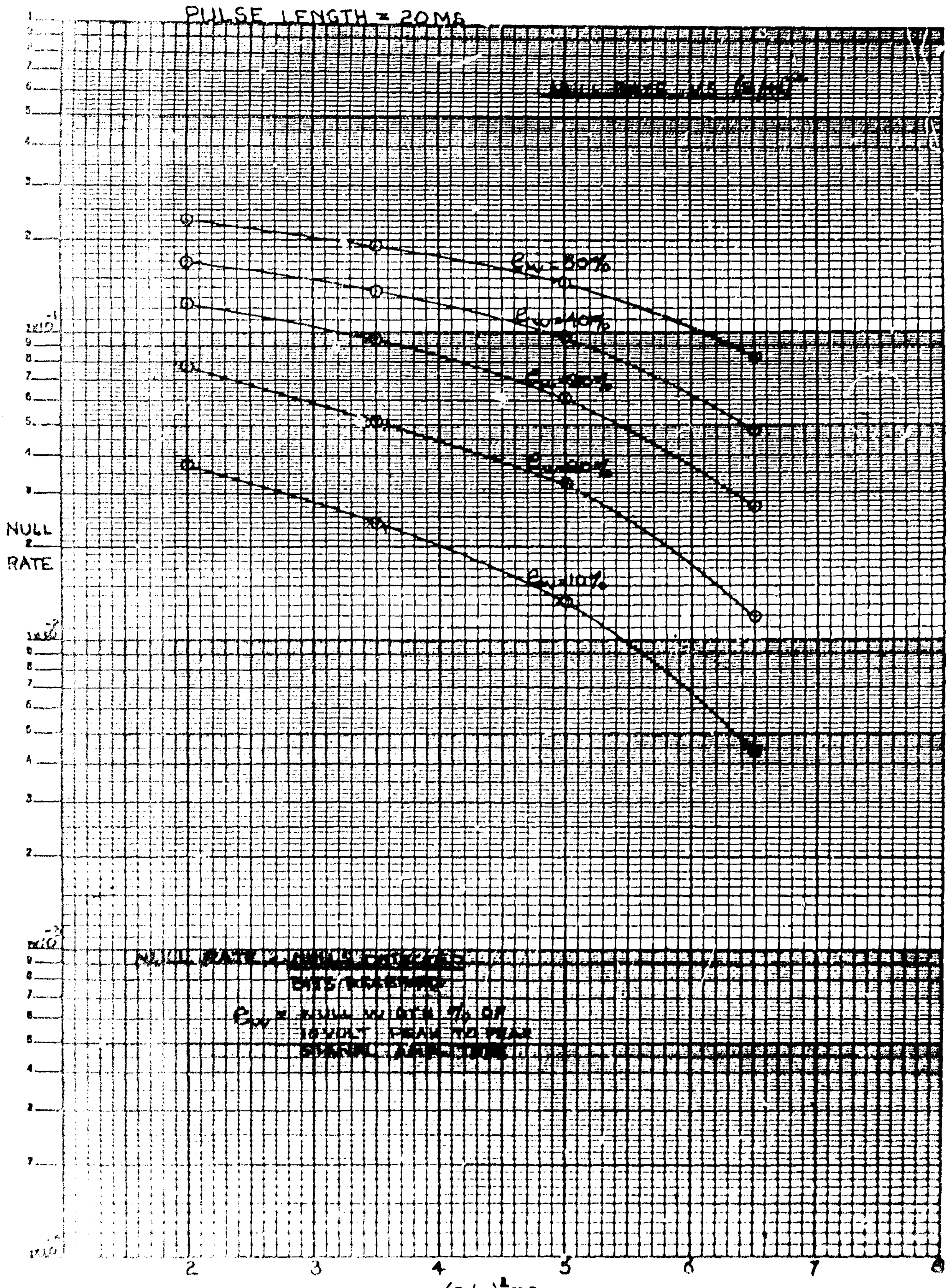

PULSE LENGTH = 20MS
NULL RATE
e_w = 50%
e_w = 40%
e_w = 30%
e_w = 20%
e_w = 10%
e_w =
2
3
4
5
6
7
8
(S/N)² DB POWER RATIO

FIG. 5.11E

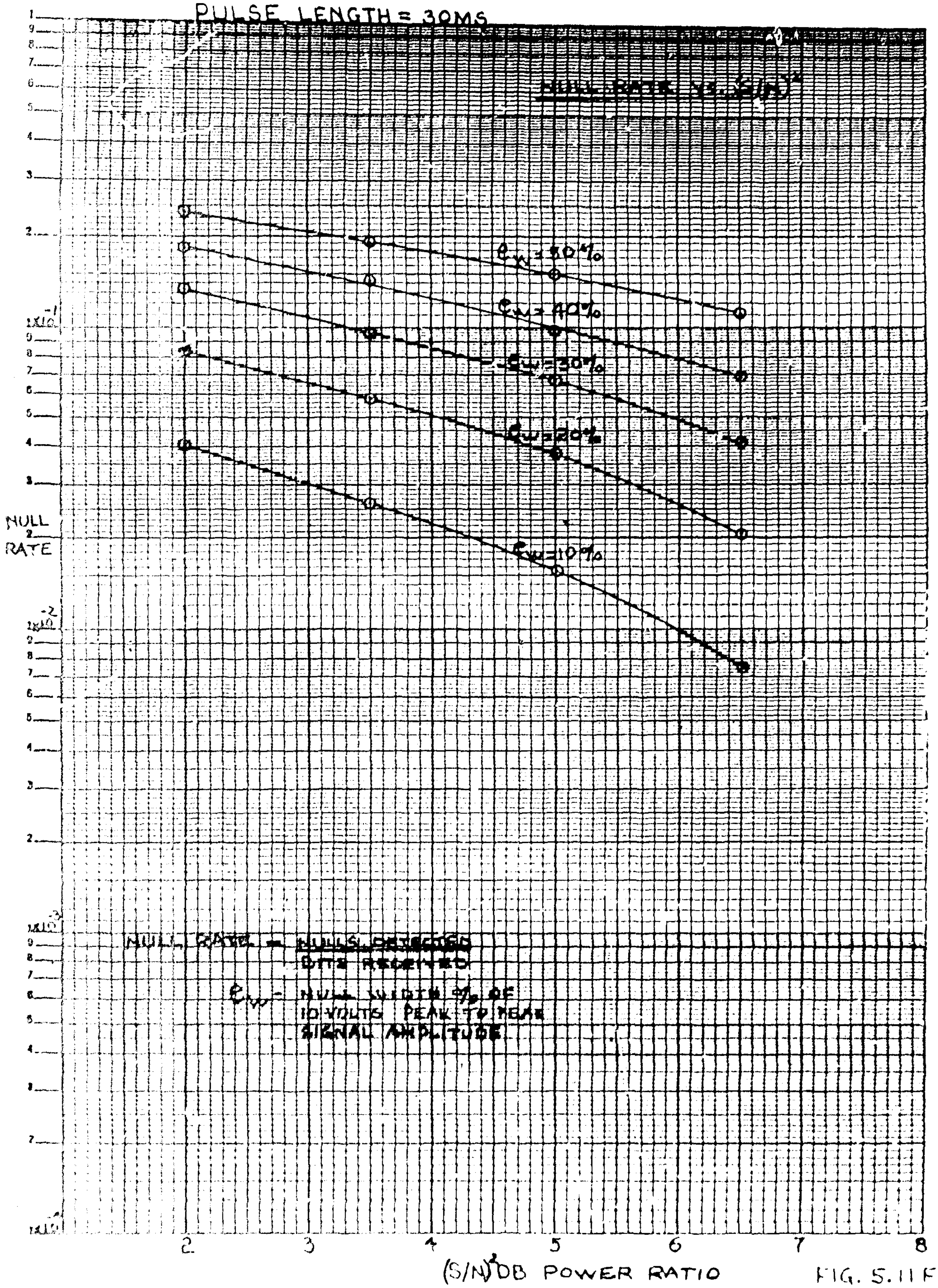
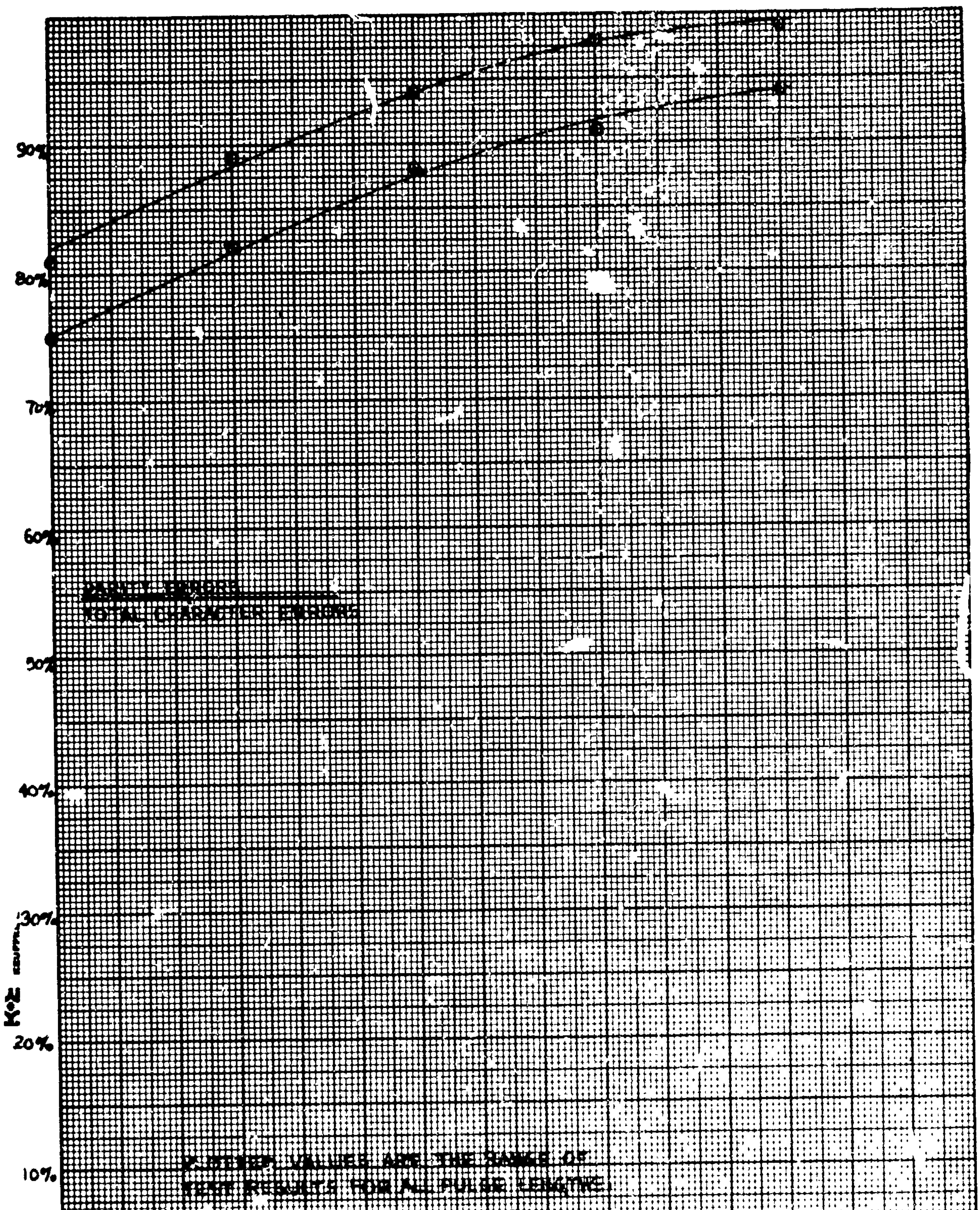

PULSE LENGTH = 30MS
NULL RATE VS. (S/N)²
e_w = 50%
e_w = 40%
e_w = 30%
e_w = 20%
e_w = 10%
NULL RATE
NULL RATE = NULLS DETECTED / BITS RECEIVED
e_w - NULL WIDTH % OF
10 VOLTS PEAK TO PEAK
SIGNAL AMPLITUDE
(S/N)² DB POWER RATIO

FIG. 5.11F



(S/N) DB POWER RATIO

FIG. 5.12A

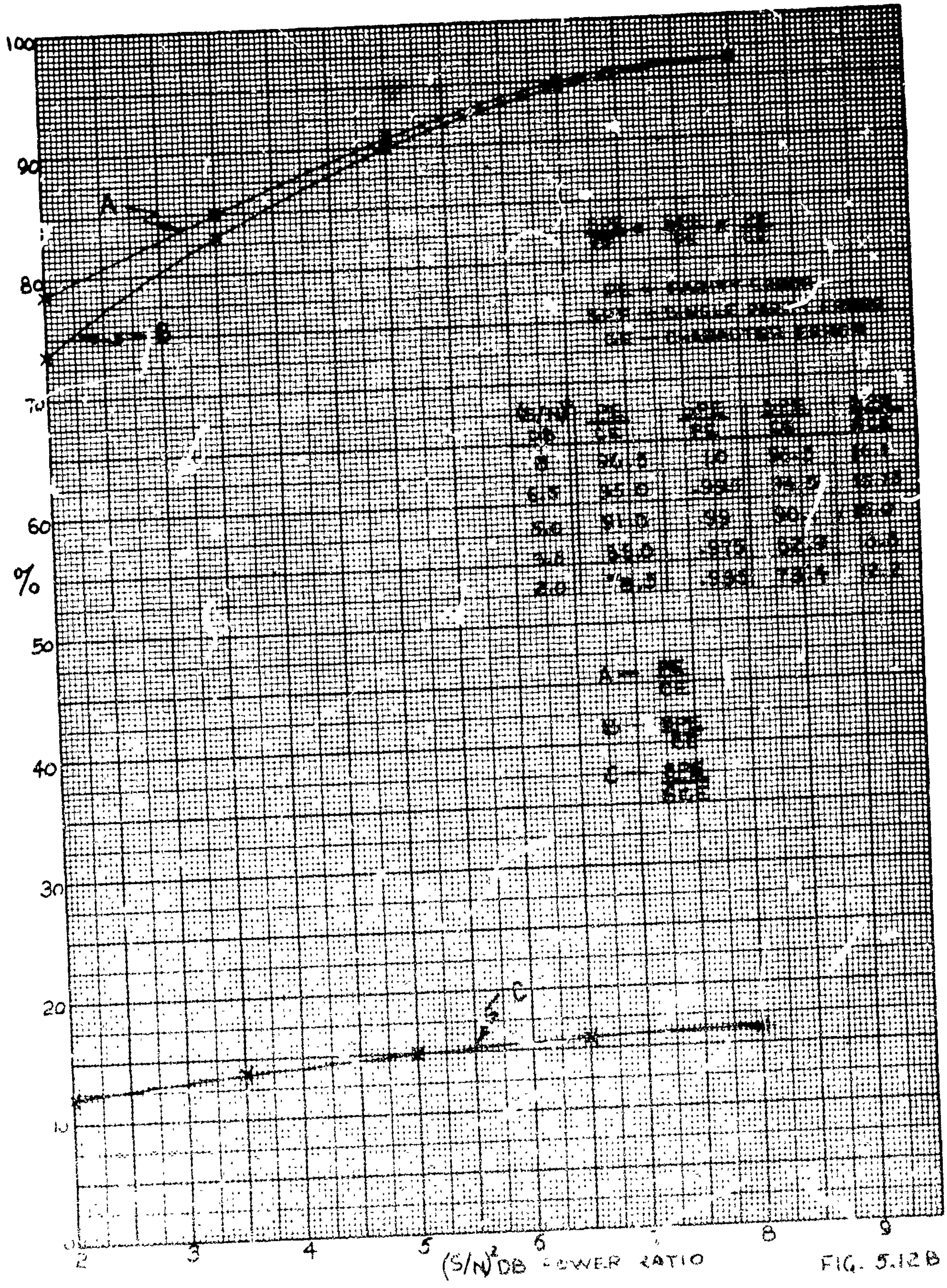
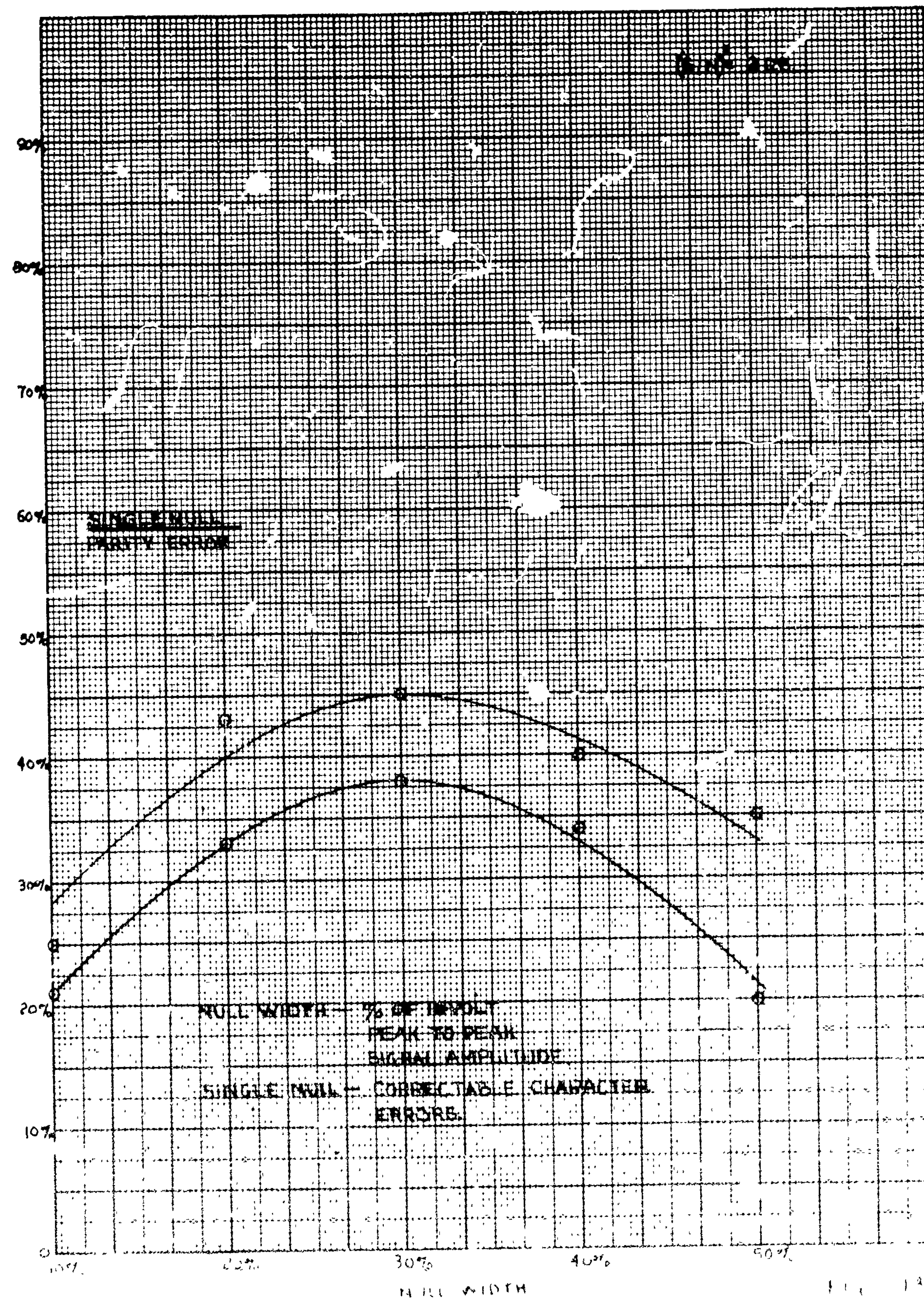
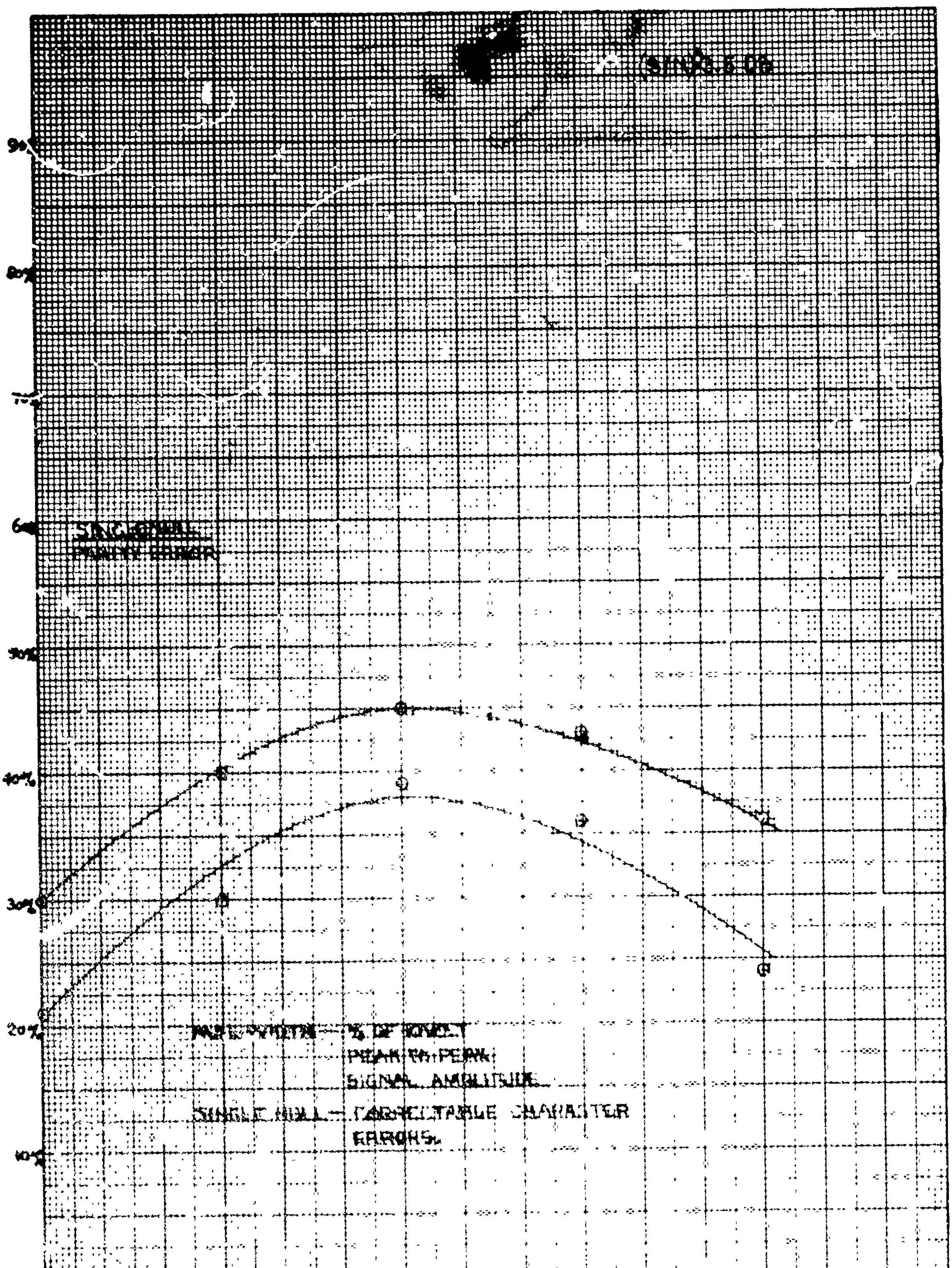


FIG. 5.12B



90%
80%
70%
60%
SINGLE NULL
PARITY ERROR
50%
40%
30%
20%
NULL WIDTH — % OF
PEAK TO PEAK
SIGNAL AMPLITUDE
SINGLE NULL — CORRECTABLE CHARACTER
ERRORS.
10%
0
30%
40%
50%
NULL WIDTH

FIG. 13A



80%
40%
30%
20%
10%
0
0%
50%
SIGNAL AMPLITUDE
ERRORS

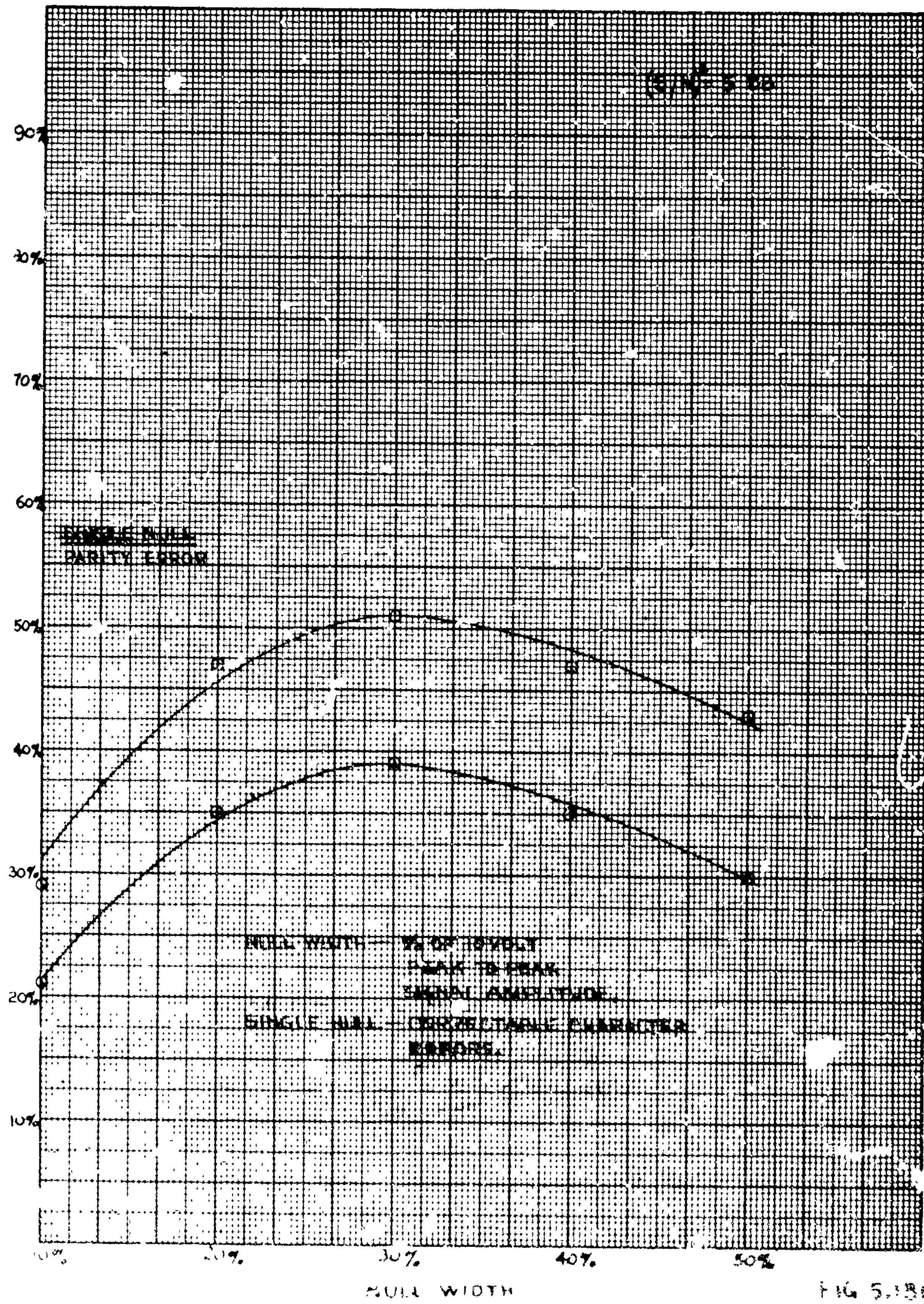


FIG 5.18C

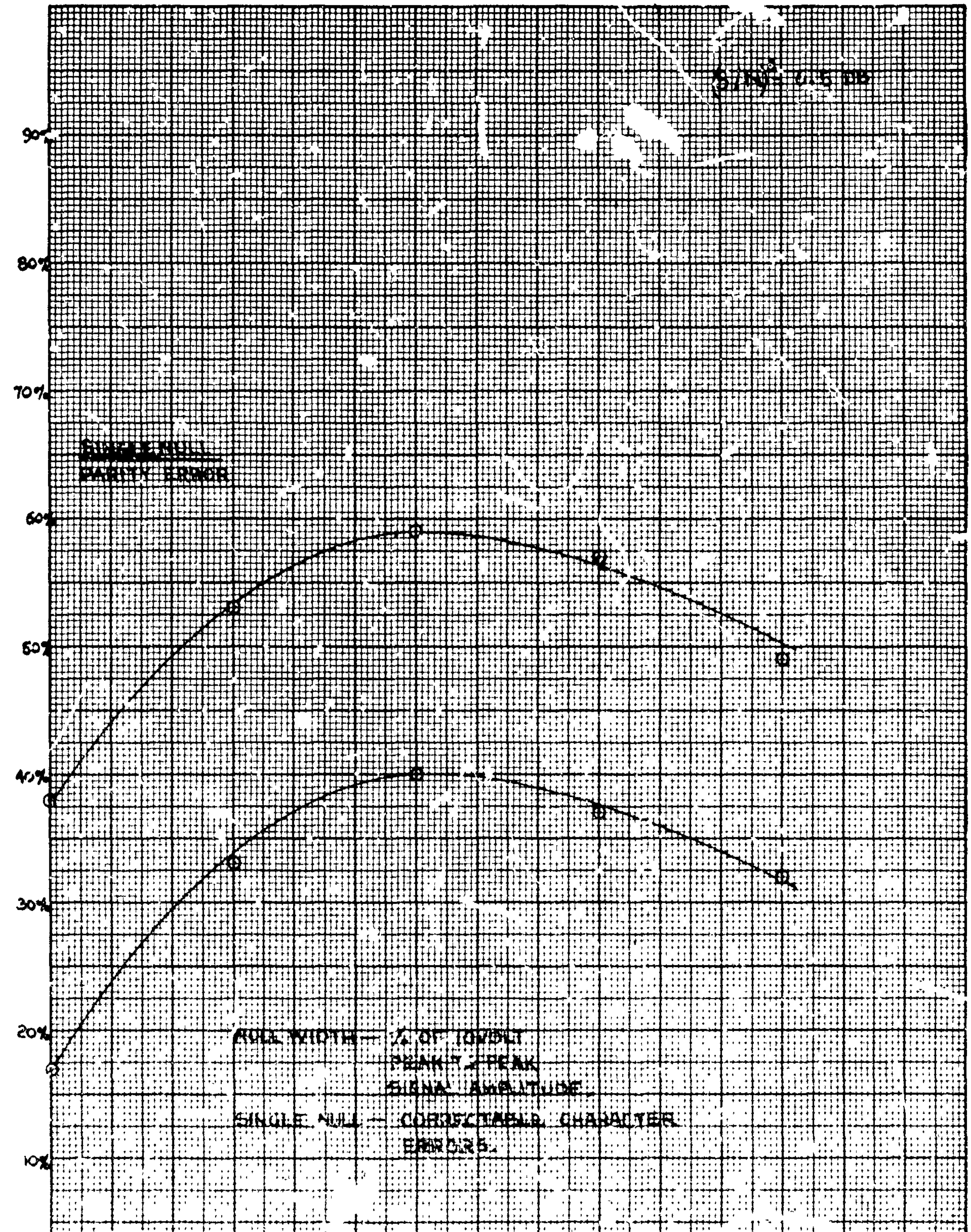
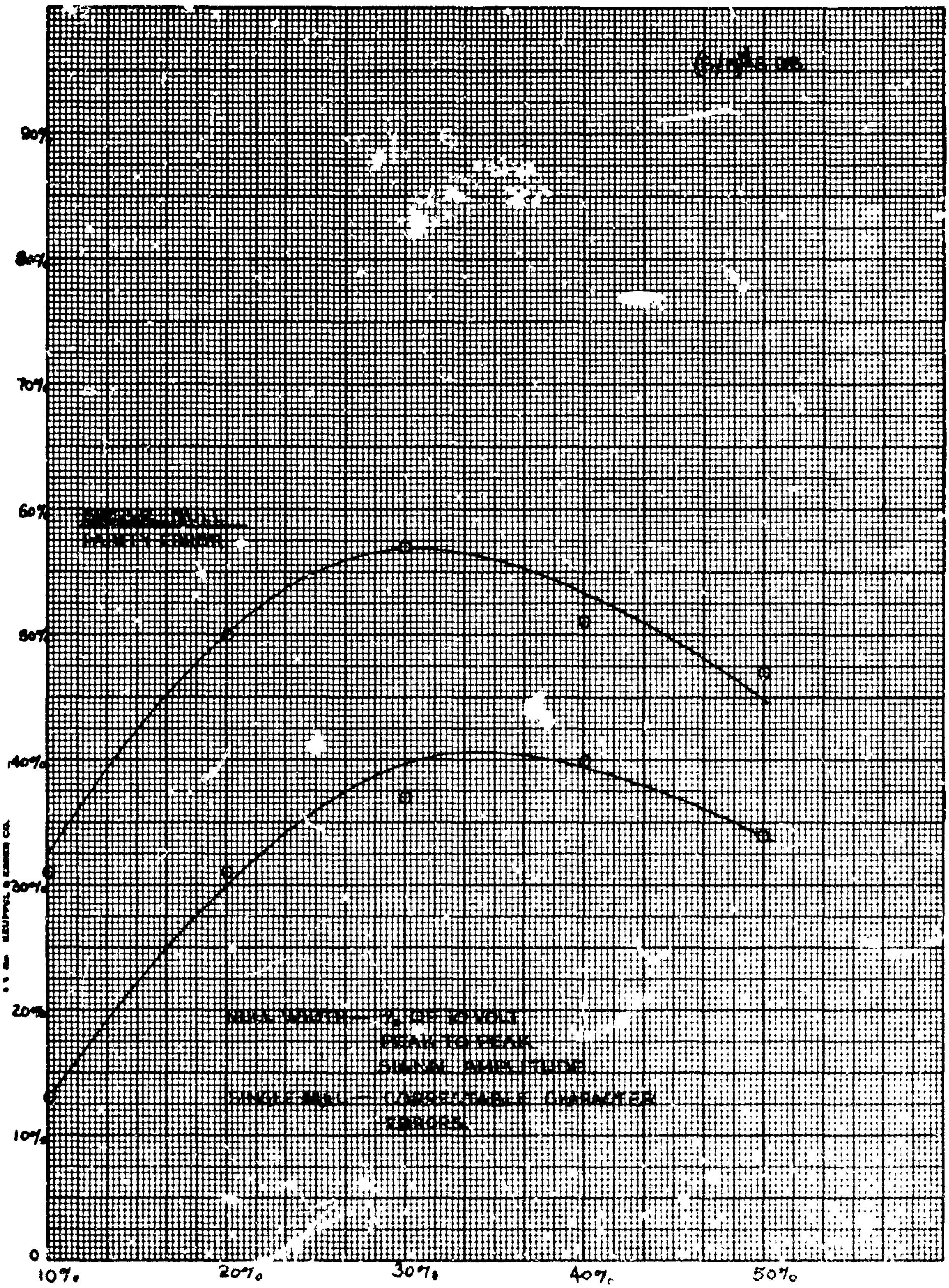

S/N = 0.5 DB
SINGLE NULL
PARITY ERROR
NULL WIDTH — % OF 10 VOLT
PEAK TO PEAK
SIGNAL AMPLITUDE.
SINGLE NULL — CORRECTABLE CHARACTER
ERRORS.
90%
80%
70%
60%
50%
40%
30%
20%
10%
0
10%
20%
30%
40%
50%
NULL WIDTH

FIG. 5.13D



90%
80%
70%
60%
50%
40%
30%
20%
10%
0
10%
20%
30%
40%
50%
PEAK TO PEAK
SIGNAL AMPLITUDE
CORRECTABLE CHARACTER
ERRORS
NULL WIDTH

FIG. 5.13E

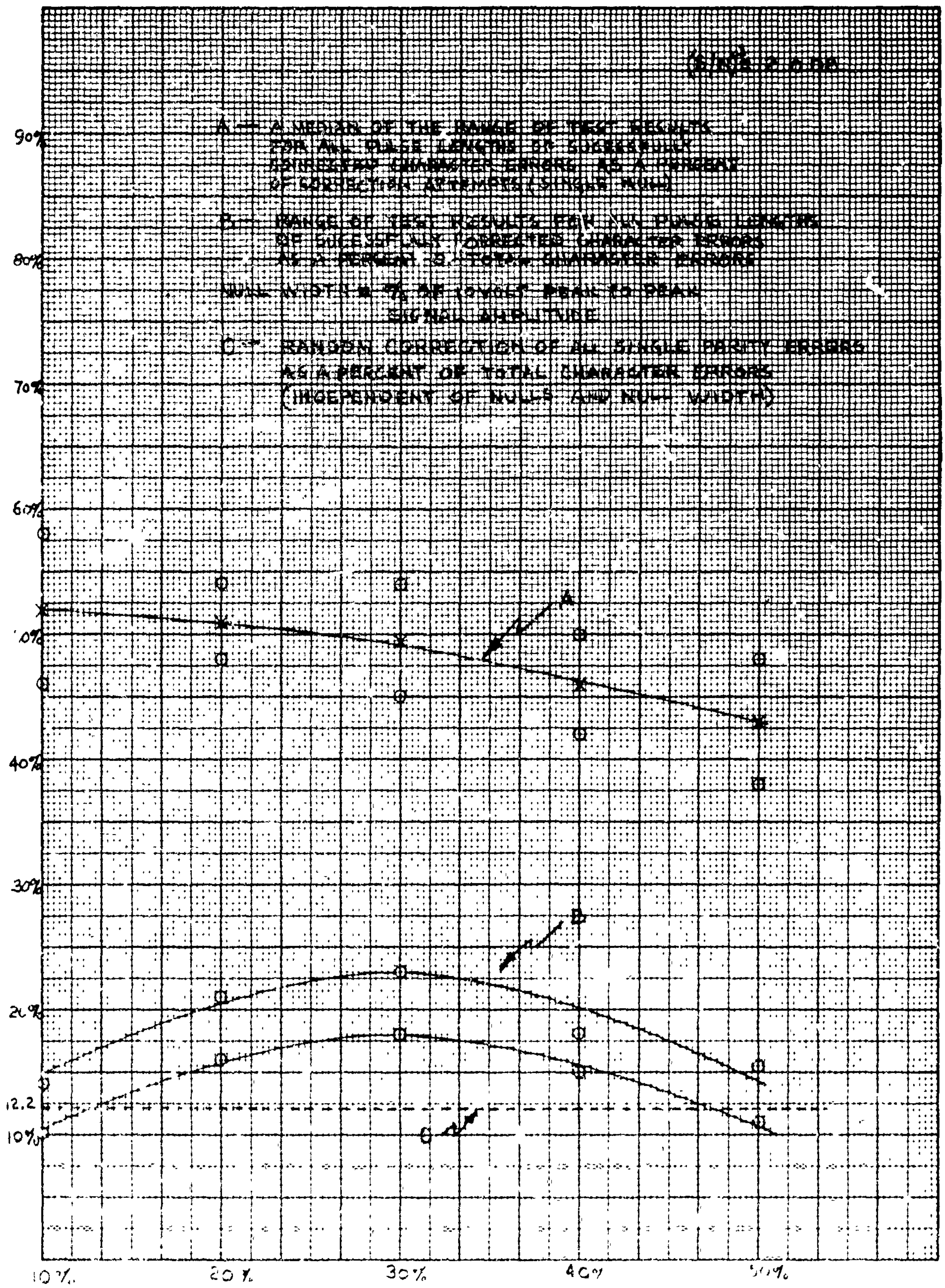
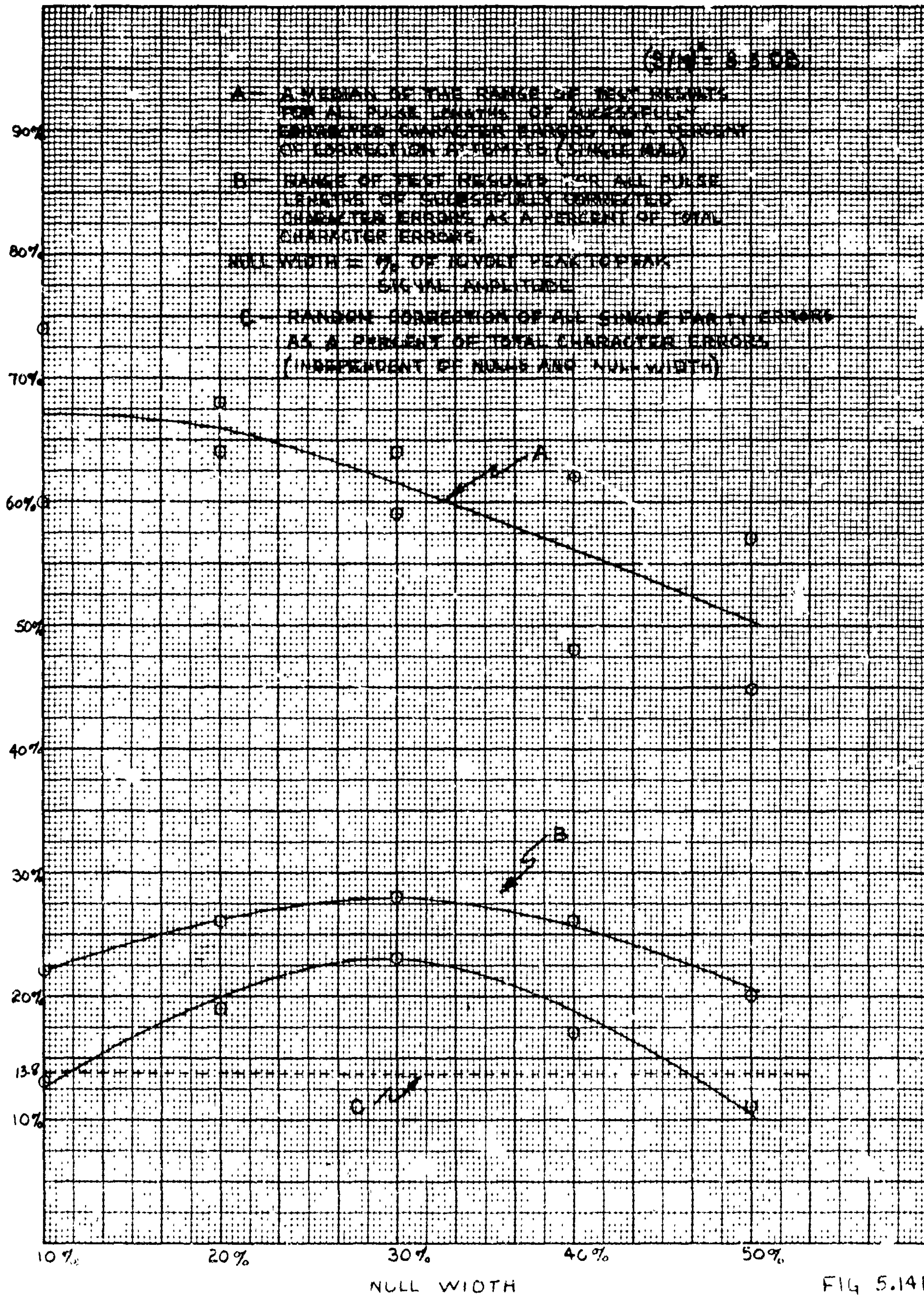

A — A MEDIAN OF THE RANGE OF TEST RESULTS FOR ALL PHASE LENGTHS OF SUCCESSFULLY CORRECTED CHARACTER ERRORS AS A PERCENT OF CORRECTION ATTEMPTS (SINGLE NULL)
B — RANGE OF TEST RESULTS FOR ALL PULSE LENGTHS OF SUCCESSFULLY CORRECTED CHARACTER ERRORS AS A PERCENT OF TOTAL CHARACTER ERRORS
NULL WIDTH = % OF 10 VOLT PEAK TO PEAK SIGNAL AMPLITUDE
C — RANDOM CORRECTION OF ALL SINGLE PARITY ERRORS AS A PERCENT OF TOTAL CHARACTER ERRORS (INDEPENDENT OF NULLS AND NULL WIDTH)
90%
80%
70%
60%
50%
40%
30%
20%
12.2
10%
A
B
C
10%
20%
30%
40%
50%
NULL WIDTH

FIG 5.14 A



B— RANGE OF TEST RESULTS FOR ALL PULSE
CHARACTER ERRORS.
C— RANDOM CORRECTION OF ALL SINGLE PARITY ERRORS
AS A PERCENT OF TOTAL CHARACTER ERRORS.
(INDEPENDENT OF NULLS AND NULL WIDTH)
90%
80%
70%
60%
50%
40%
30%
20%
13.8
10%
A
B
C
10%
20%
30%
40%
50%
NULL WIDTH

FIG 5.14B

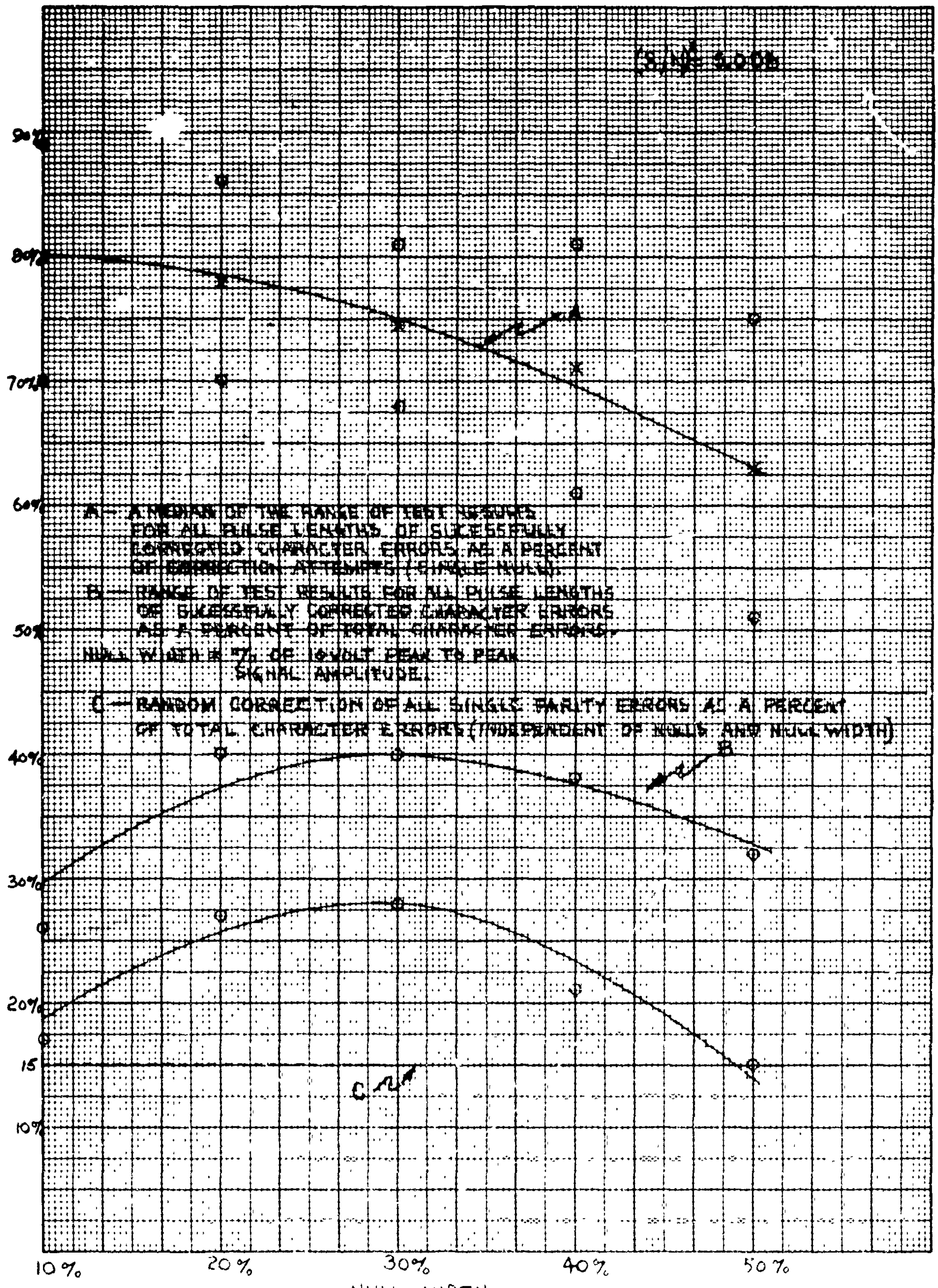
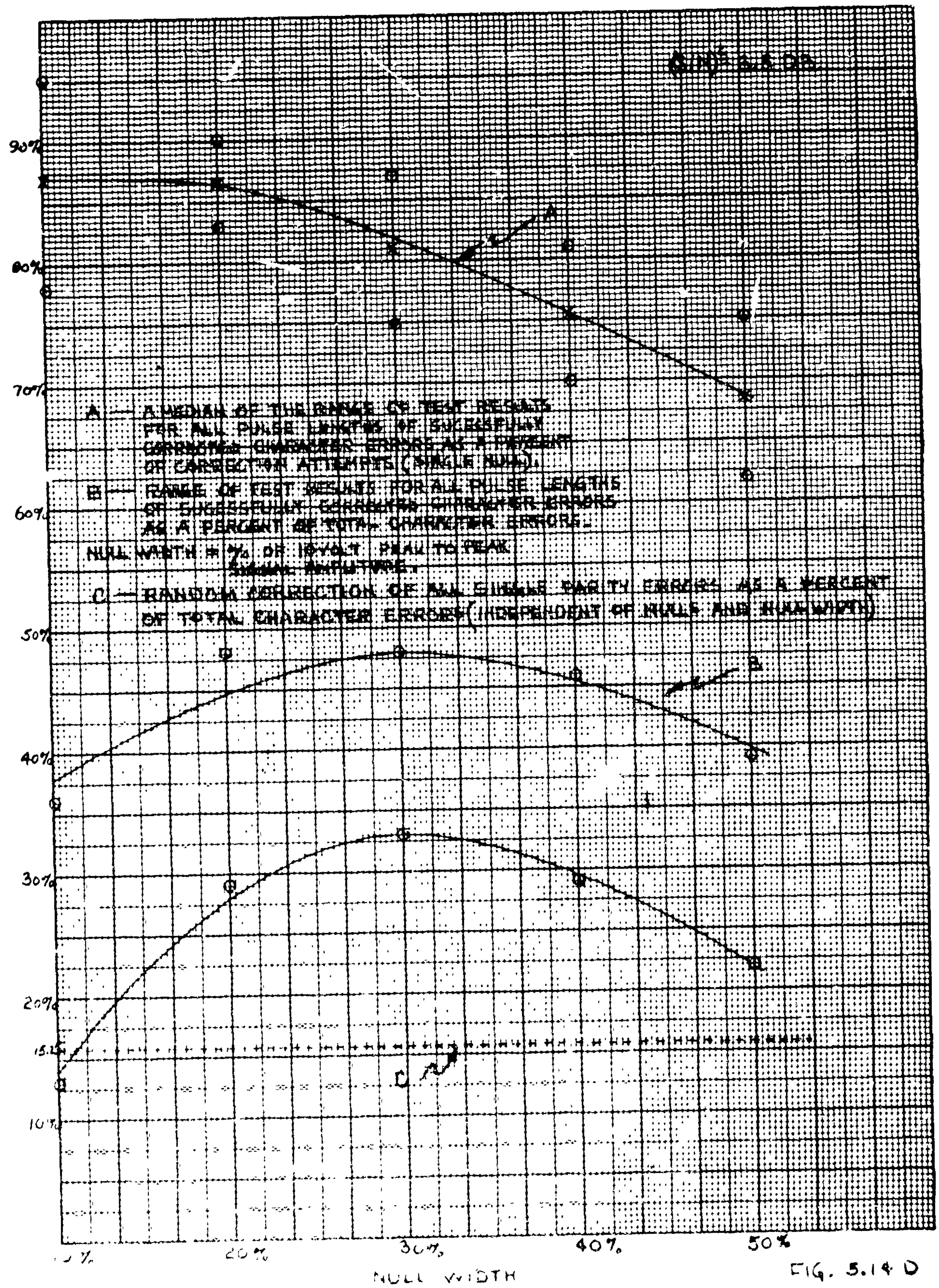

90%
80%
70%
60%
50%
40%
30%
20%
15
10%
A — A MEDIAN OF THE RANGE OF TEST RESULTS FOR ALL PULSE LENGTHS OF SUCESSFULLY CORRECTED CHARACTER ERRORS AS A PERCENT OF CORRECTION ATTEMPTS (SINGLE NULL).
B — RANGE OF TEST RESULTS FOR ALL PULSE LENGTHS OF SUCESSFULLY CORRECTED CHARACTER ERRORS AS A PERCENT OF TOTAL CHARACTER ERRORS.
NULL WIDTH IN % OF 10 VOLT PEAK TO PEAK SIGNAL AMPLITUDE.
C — RANDOM CORRECTION OF ALL SINGLE PARITY ERRORS AS A PERCENT OF TOTAL CHARACTER ERRORS (INDEPENDENT OF NULLS AND NULL WIDTH)
A
B
C
10%
20%
30%
40%
50%
NULL WIDTH

FIG. 5.146



90%
80%
70%
60%
50%
40%
30%
20%
10%
A
A — A MEDIAN OF THE RANGE OF TEST RESULTS FOR ALL PULSE LENGTHS OF SUCCESSFULLY CORRECTED CHARACTER ERRORS AS A PERCENT OF CORRECTION ATTEMPTS (SINGLE NULL).
B — RANGE OF TEST RESULTS FOR ALL PULSE LENGTHS OF SUCCESSFULLY CORRECTED CHARACTER ERRORS AS A PERCENT OF TOTAL CHARACTER ERRORS.
NULL WIDTH = % OF 10 VOLT PEAK TO PEAK SIGNAL AMPLITUDE.
C — RANDOM CORRECTION OF ALL SINGLE PARITY ERRORS AS A PERCENT OF TOTAL CHARACTER ERRORS (INDEPENDENT OF NULLS AND NULL WIDTH)
B
C
20%
30%
40%
50%
NULL WIDTH

FIG. 5.14 D

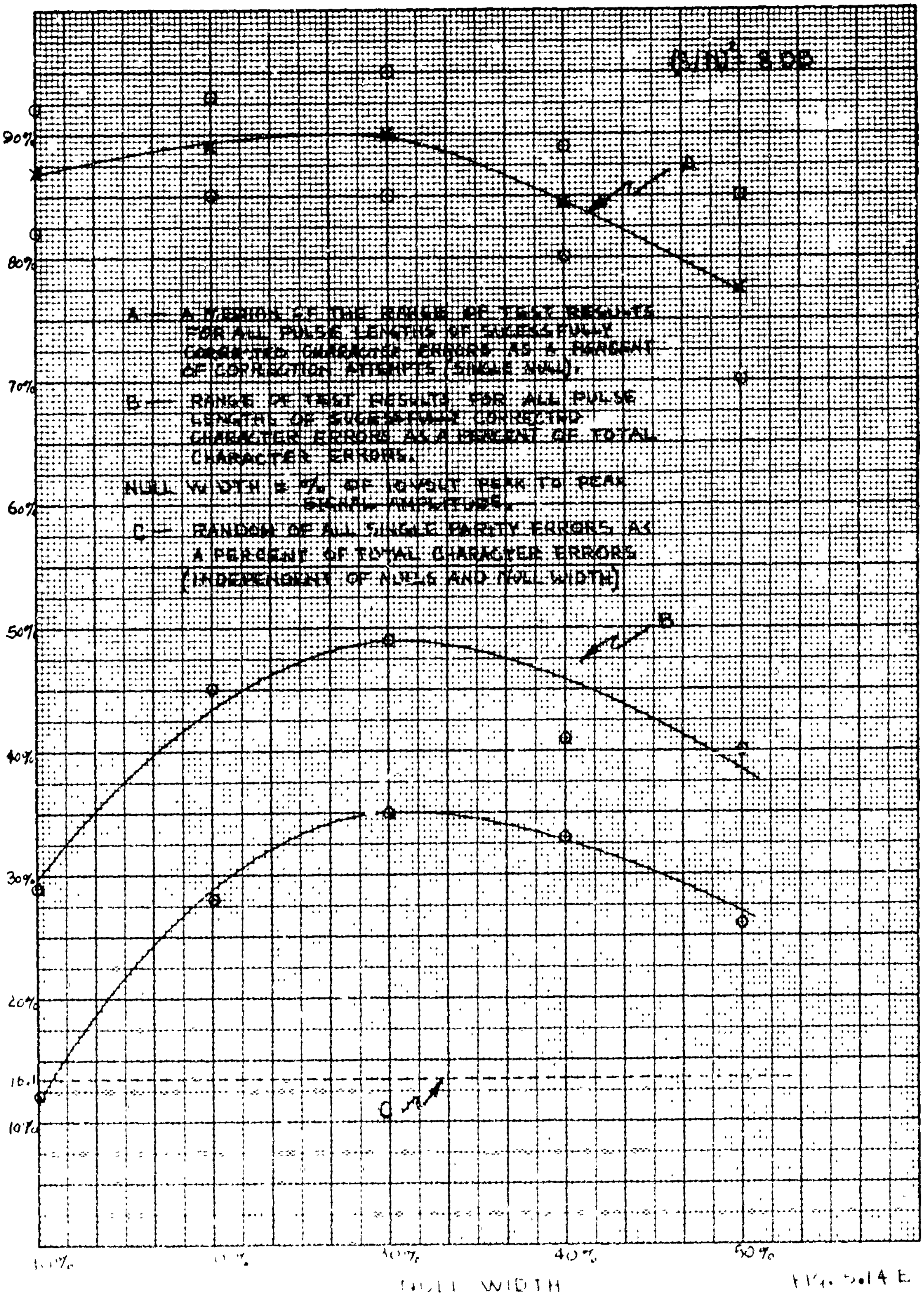
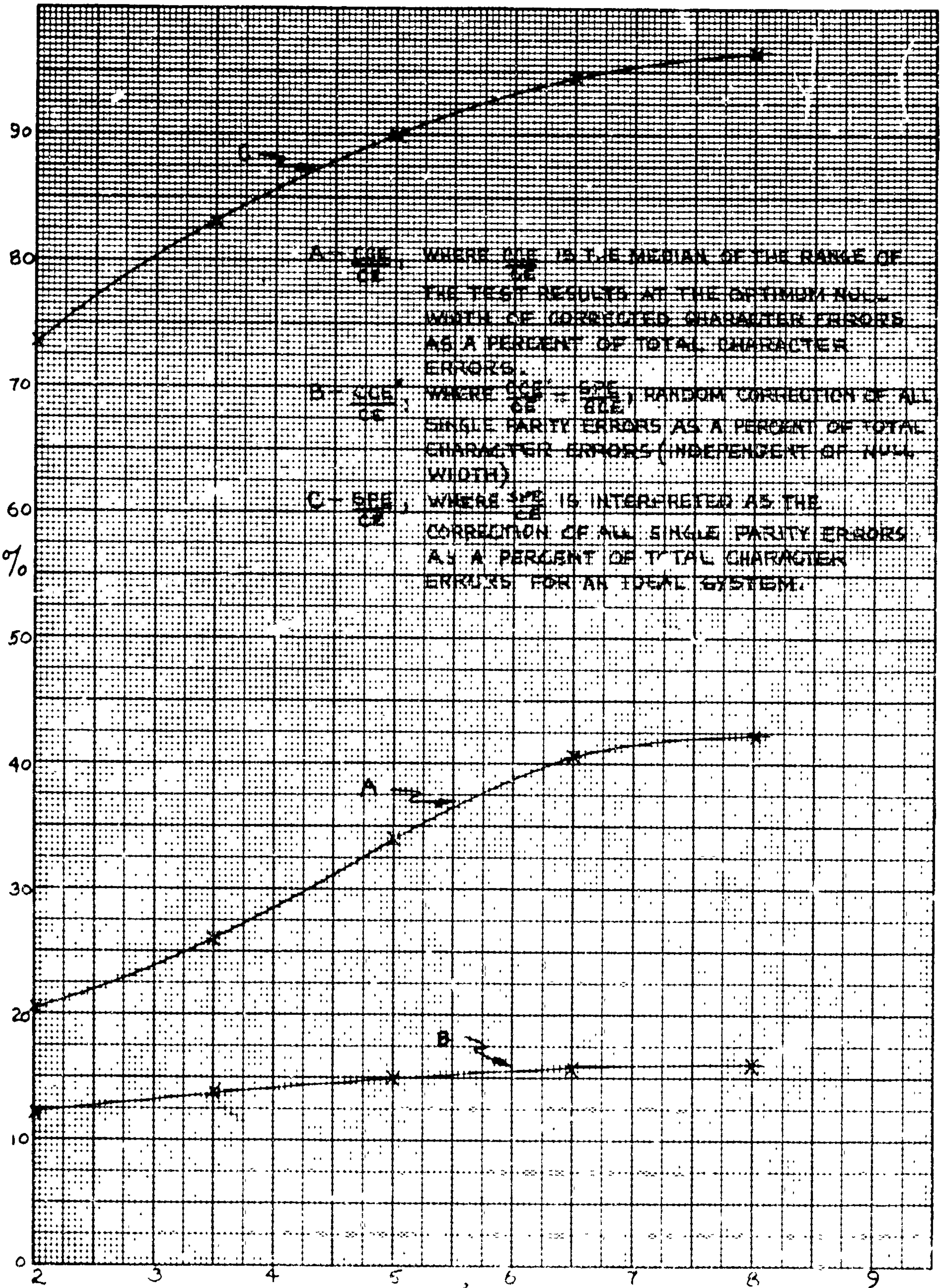

90%
80%
70%
60%
50%
40%
30%
20%
16.1
10%
A
B
C
FOR ALL PULSE LENGTHS OF SUCCESSFULLY
OF CORRECTION ATTEMPTS (SINGLE NULL).
B — RANGE OF TEST RESULTS FOR ALL PULSE
CHARACTER ERRORS.
C — RANDOM OF ALL SINGLE PARITY ERRORS AS
A PERCENT OF TOTAL CHARACTER ERRORS
(INDEPENDENT OF NULLS AND NULL WIDTH)
40%
50%
NULL WIDTH

FIG. 5.14 E



90
80
70
60
%
50
40
30
20
10
0
2
3
4
5
6
7
8
9
C
A
B
A = CCE/CE, WHERE CCE/CE IS THE MEDIAN OF THE RANGE OF THE TEST RESULTS AT THE OPTIMUM NULL WIDTH OF CORRECTED CHARACTER ERRORS AS A PERCENT OF TOTAL CHARACTER ERRORS.
B = CCE/CE, WHERE CCE/CE = SPE/CCE, RANDOM CORRECTION OF ALL SINGLE PARITY ERRORS AS A PERCENT OF TOTAL CHARACTER ERRORS (INDEPENDENT OF NULL WIDTH)
C = SPE/CE, WHERE SPE/CE IS INTERPRETED AS THE CORRECTION OF ALL SINGLE PARITY ERRORS AS A PERCENT OF TOTAL CHARACTER ERRORS FOR AN IDEAL SYSTEM.
$(S/N)^2$ DA WER RATIO

FIG 5.14F

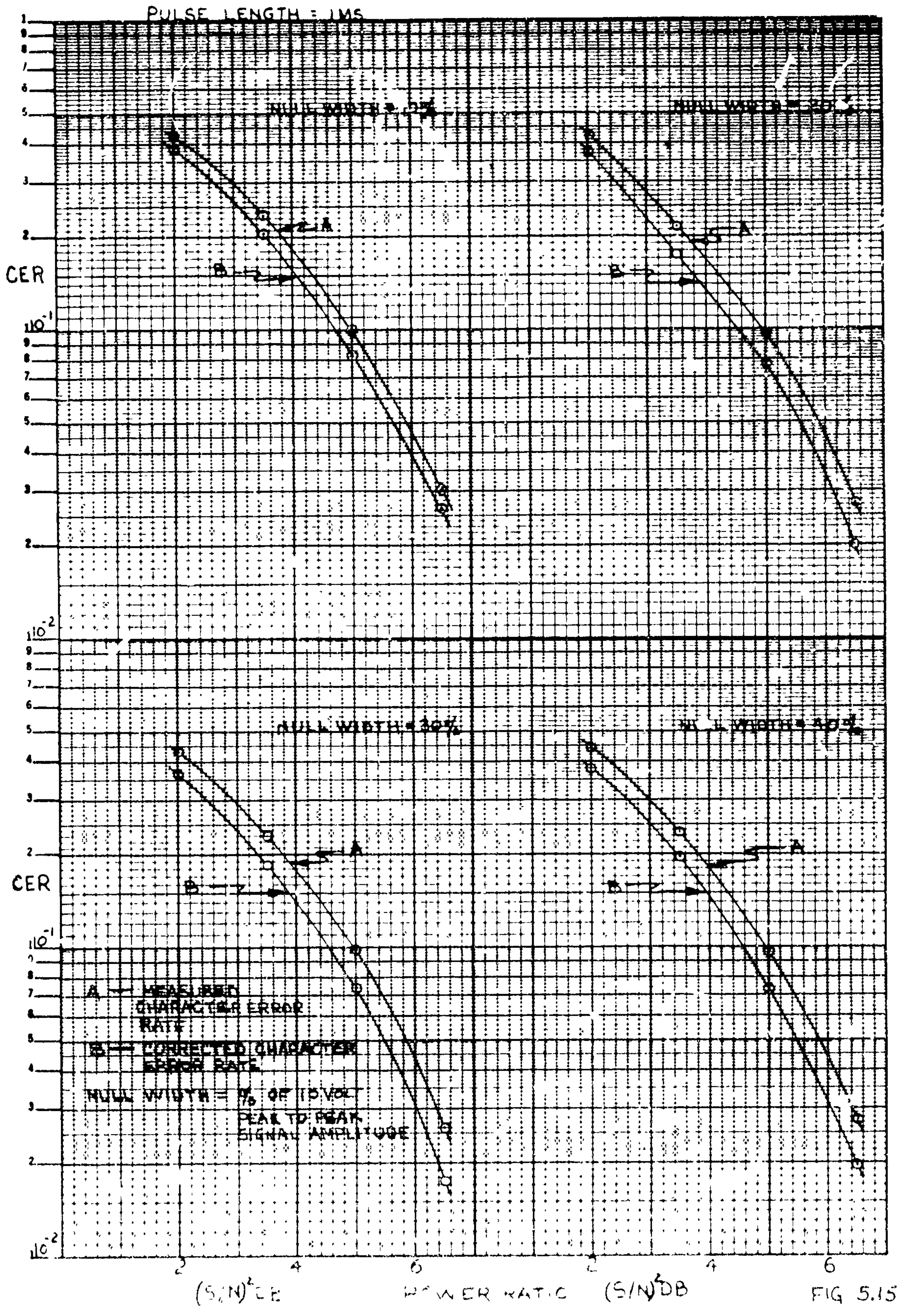
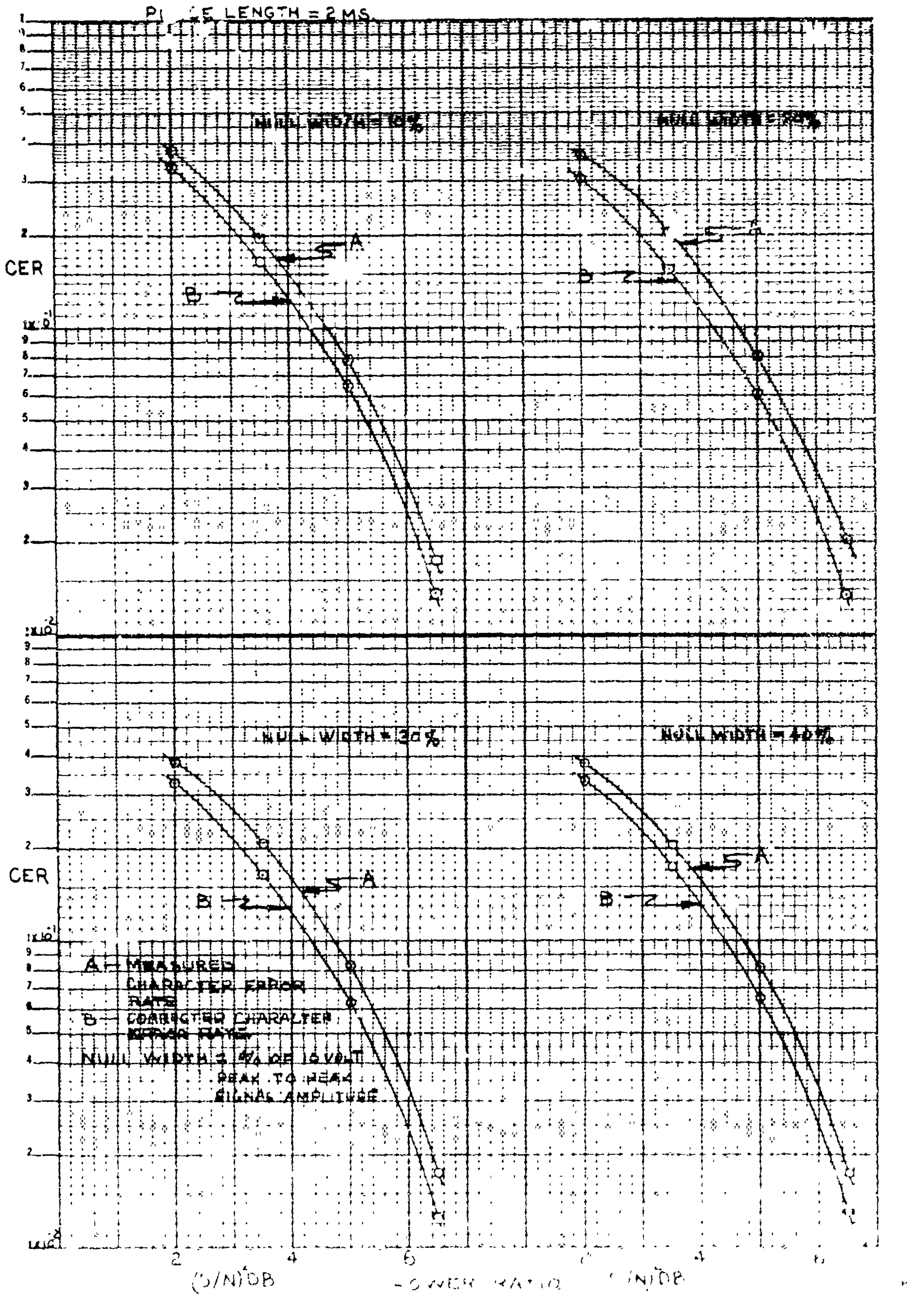
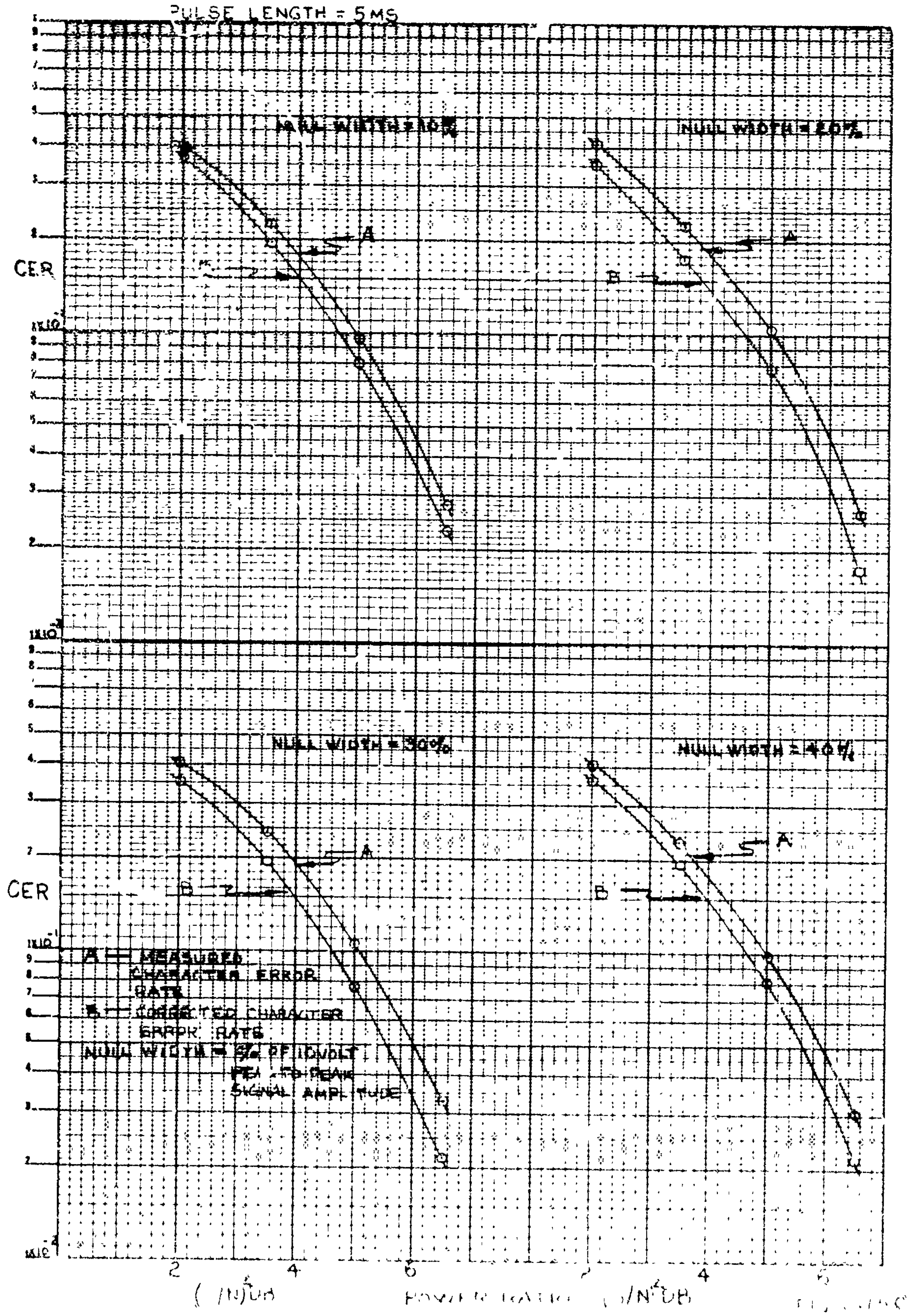
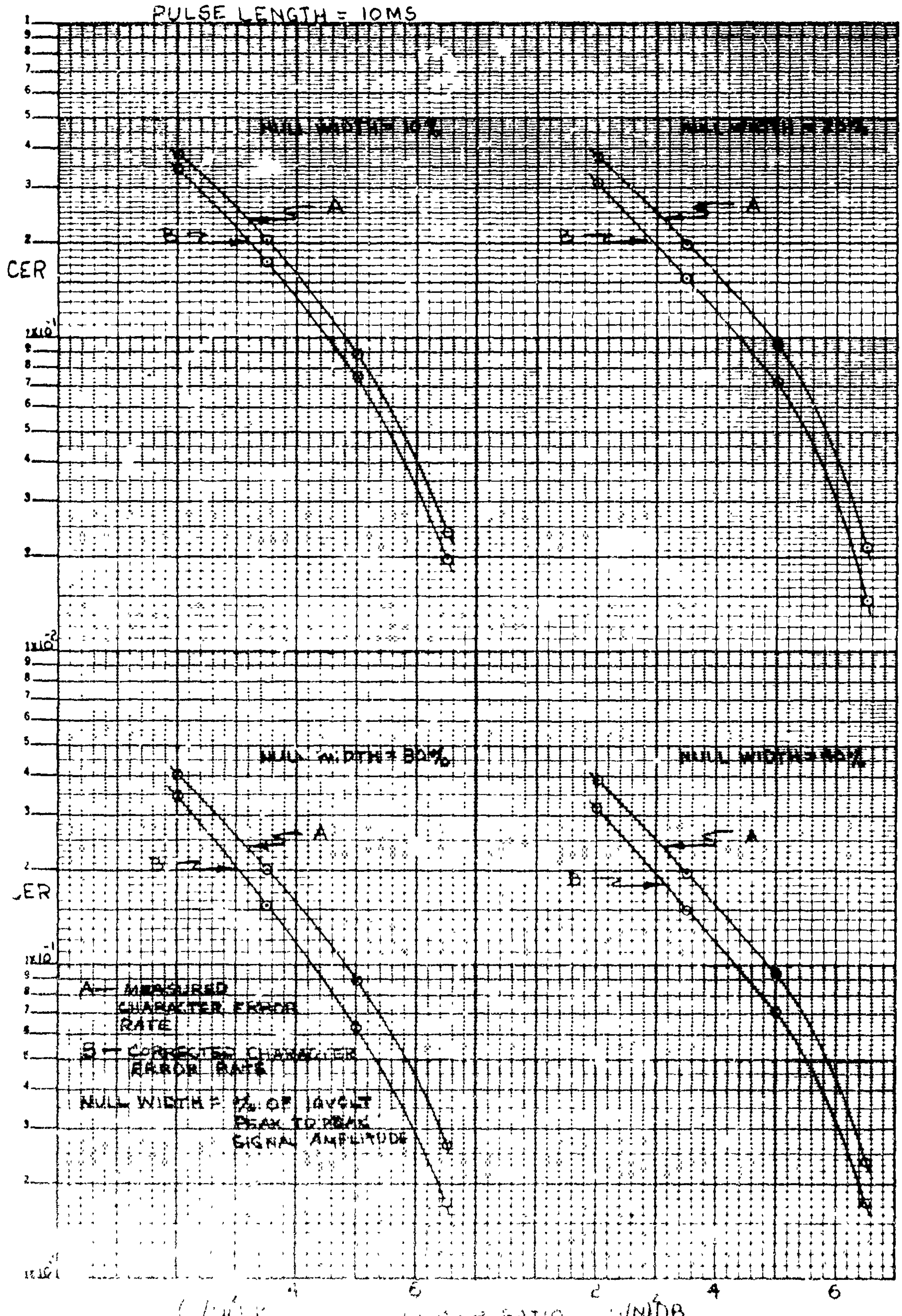

PULSE LENGTH = 1MS
CER
NULL WIDTH = 30%
A — MEASURED CHARACTER ERROR RATE
B — CORRECTED CHARACTER ERROR RATE
NULL WIDTH = % OF 15 VOLT PEAK TO PEAK SIGNAL AMPLITUDE
A
B
POWER RATIO
(S/N)² DB

FIG 5.15A

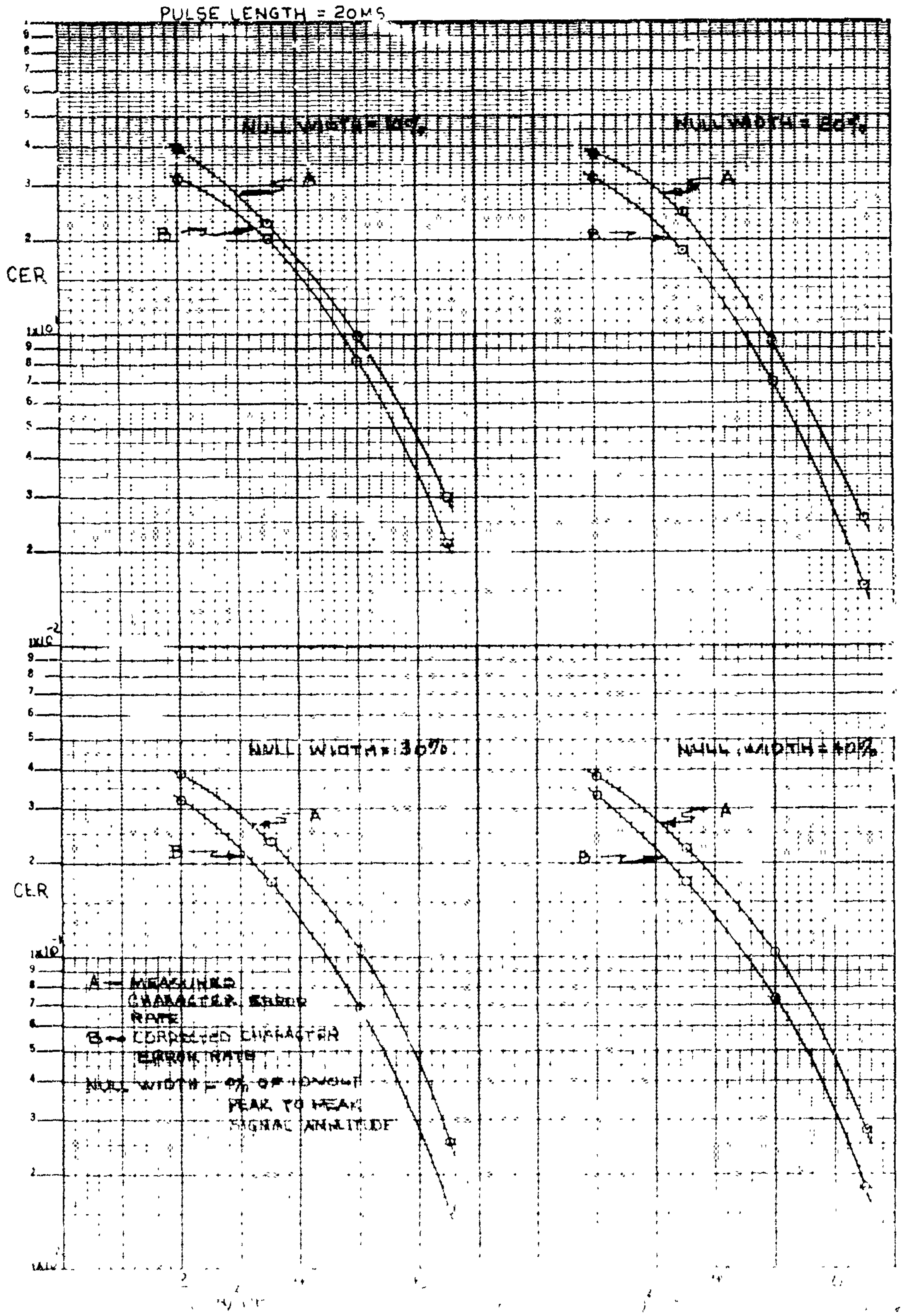

PULSE LENGTH = 2 MS
NULL WIDTH = 10%
NULL WIDTH = 20%
NULL WIDTH = 30%
NULL WIDTH = 40%
CER
CER
A
B
A — MEASURED CHARACTER ERROR RATE
B — CORRECTED CHARACTER ERROR RATE
NULL WIDTH = % OF 10 VOLT PEAK TO PEAK SIGNAL AMPLITUDE
(S/N)DB
POWER RATIO
(S/N)DB

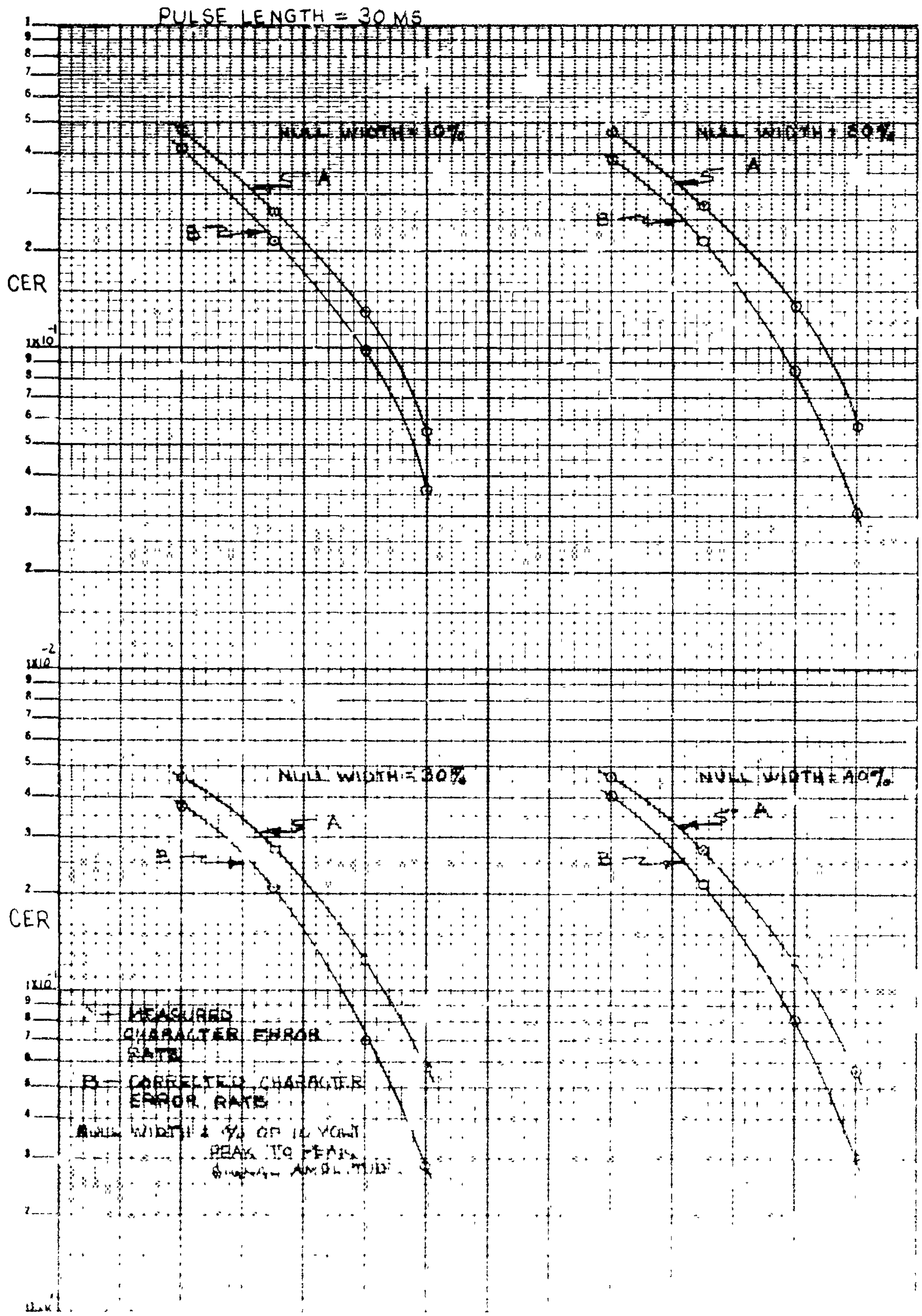

PULSE LENGTH = 5MS
NULL WIDTH = 10%
NULL WIDTH = 20%
NULL WIDTH = 30%
NULL WIDTH = 40%
CER
CER
A
B
A — MEASURED CHARACTER ERROR RATE
B — CORRECTED CHARACTER ERROR RATE
NULL WIDTH = % OF IDUDLT PEAK TO PEAK SIGNAL AMPLITUDE
2
4
6
4

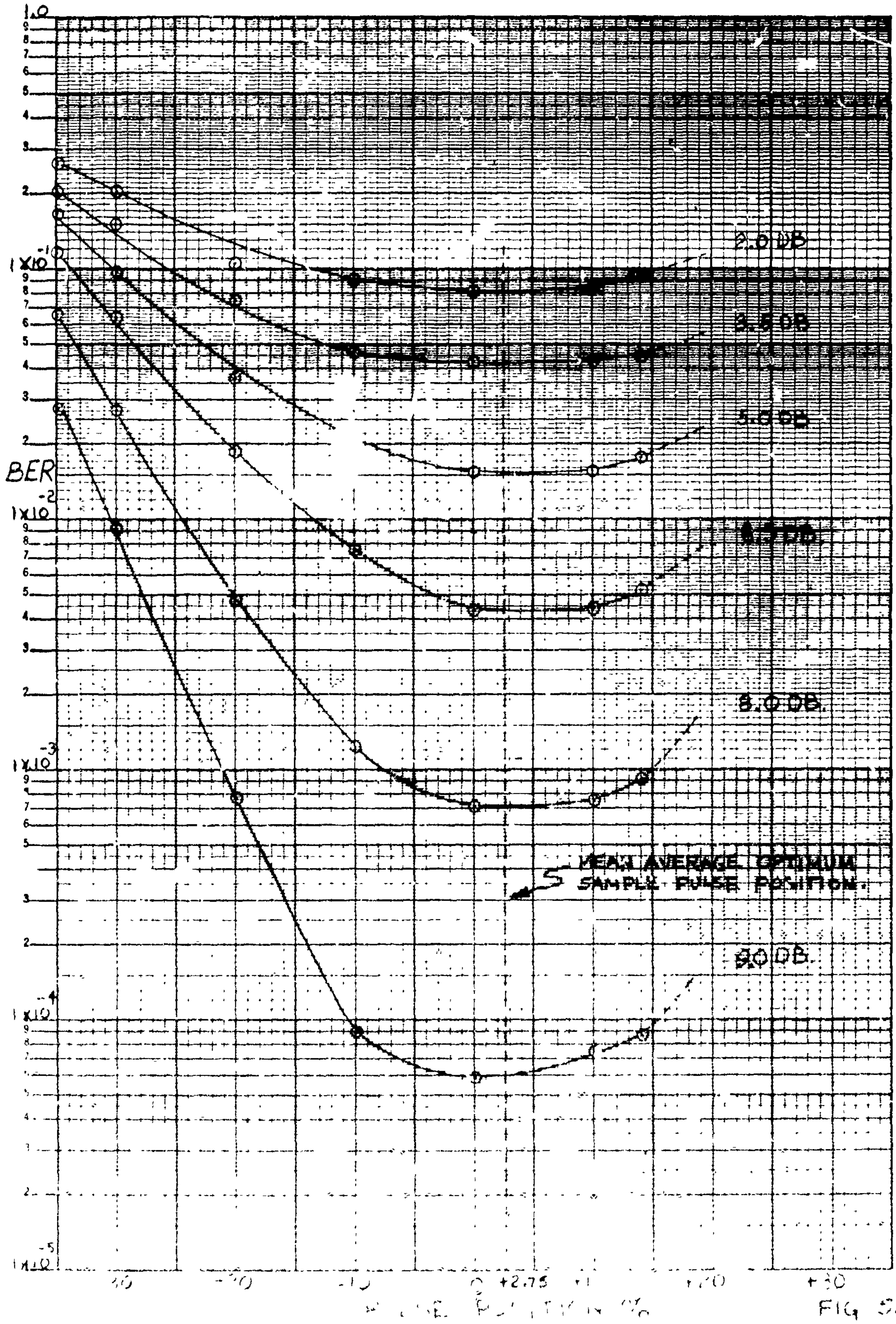


PULSE LENGTH = 10MS
NULL WIDTH 10%
NULL WIDTH = 20%
NULL WIDTH = 30%
NULL WIDTH = 40%
CER
A
B
A — MEASURED CHARACTER ERROR RATE
B — CORRECTED CHARACTER ERROR RATE
NULL WIDTH = % OF 10VOLT PEAK TO PEAK SIGNAL AMPLITUDE
(S/N)DB

FIG 5.


PULSE LENGTH = 20MS
NULL WIDTH = 20%
CER
NULL WIDTH = 30%
NULL WIDTH = 40%
CER
A — MEASURED CHARACTER ERROR RATE
B — CORRECTED CHARACTER ERROR RATE
PEAK TO PEAK SIGNAL AMPLITUDE


PULSE LENGTH = 30 MS
NULL WIDTH = 10%
NULL WIDTH = 20%
NULL WIDTH = 30%
NULL WIDTH = 40%
CER
A
B
A — MEASURED CHARACTER ERROR RATE
B — CORRECTED CHARACTER ERROR RATE
NULL WIDTH = % OF 12 VOLT PEAK TO PEAK SIGNAL AMPLITUDE



BER
2.0 DB
3.5 DB
5.0 DB
6.5 DB
8.0 DB
9.0 DB
MEAN AVERAGE OPTIMUM SAMPLE PULSE POSITION
+2.75
+30

FIG 5.16

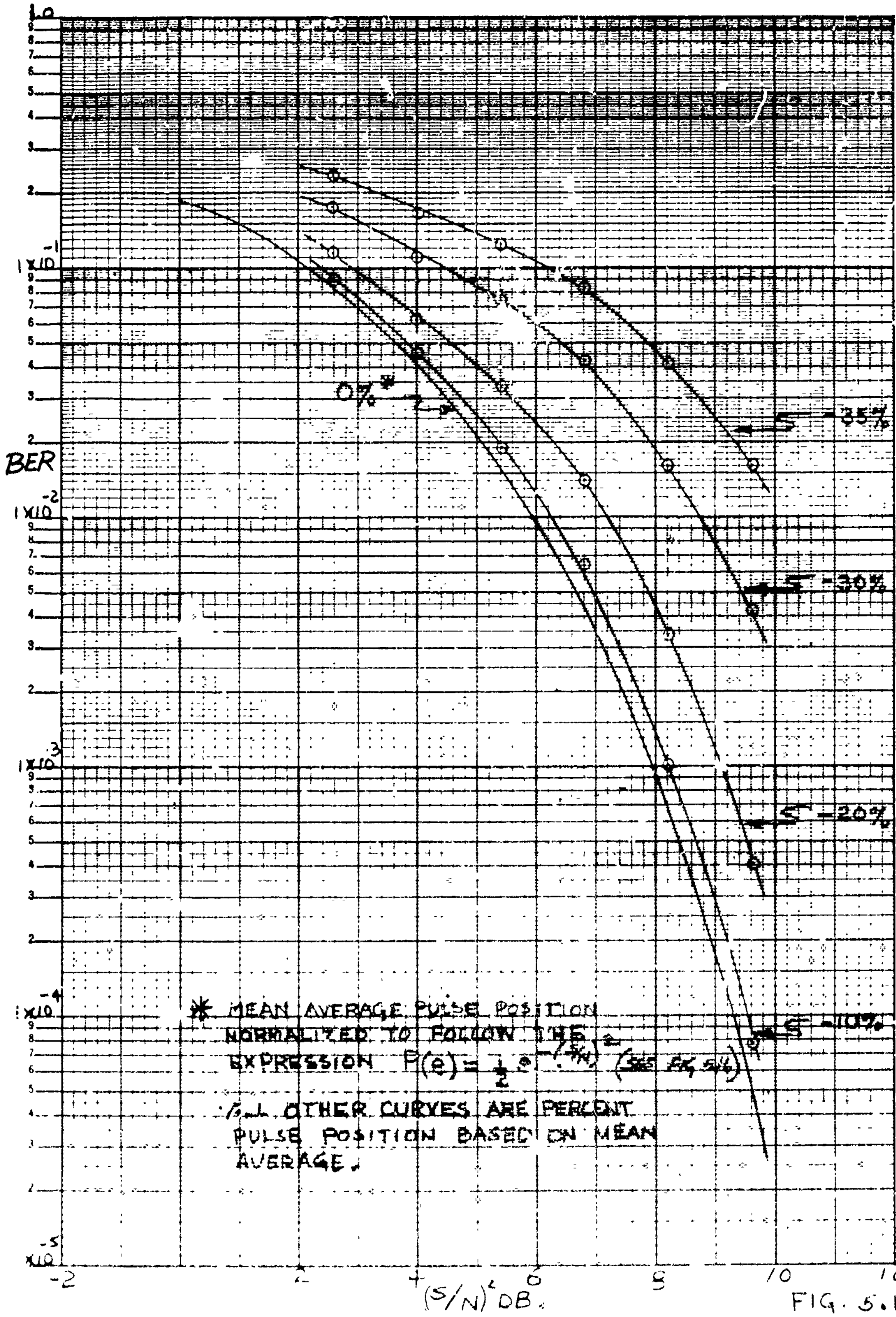

BER
(S/N)² DB
0%*
δ = 35%
δ = 30%
δ = 20%
δ = 10%
* MEAN AVERAGE PULSE POSITION NORMALIZED TO FOLLOW THE EXPRESSION P(e) = ½ e^-(S/N)² (SEE FIG 5.16)
ALL OTHER CURVES ARE PERCENT PULSE POSITION BASED ON MEAN AVERAGE.

FIG. 5.17

SECTION 6.0 SOME SYSTEM CONSIDERATIONS

6.1 Introduction

The problems of system design are certainly complex and the methods are not formalized. Experience, good judgement, compromise, evaluation of alternatives - all are factors in arriving at the final design, and although every reasonable effort may be expended to maximize the probability of success, the system designer seldom achieves that inner security which comes from the knowledge that he has arrived at the best possible configuration.

At the present time, the communication requirement for air traffic control is not completely known, and the dimensions and ramifications of the problem are so widespread and diverse that to embody all of them in a single unified concept may well be impossible. However, it is always possible to examine different alternatives in the light of practical limitations and known requirements. Ideally, it should also be possible to assign weighting factors to each and so arrive at a figure of merit; however, this final step is the province of the system engineer, and no attempt will be made here to infringe upon his difficult art.

The following discussion presents some alternatives, discusses certain pros and cons, and lists various practical considerations.

6.2 Sequencing

Consider the timing diagram of Fig. 6.1 as a starting point. Here, two methods of sequencing two-way communication systems are shown. In 6.1 (A), a ground station transmits, waits for a reply, and so on, so that none of the transmissions in the system overlap in time. In 6.1 (B) the overlapping restriction is not applied; continuous ground-to-air transmissions are permitted and air-to-ground transmissions from different aircraft may follow each other closely with very little time separation. However, the starting time of each transmission is subject to certain limitations which appear to be necessary to prevent interference between air-to-ground messages. The beginning of ground-to-air messages follows either the end of the last ground-to-air message, or the end of the last air-to-ground response, whichever is later; and the beginning of an air-to-ground message coincides approximately with the beginning of a ground-to-air message. These restrictions result in occasional gaps between signals, as shown in 6.1 (B).

For convenience in what follows, these two different approaches will be referred to as sequencing Plan A or B, or simply Plan A or B.

6.3 Communication Channel Requirements

6.3.1 General

Networks operating according to the sequencing plans A or B can be operated on any duplex communication facility,

such as four wire telephone circuits or dual channel radio. However, A is the only plan usable on a two-wire or single-channel circuit. It is not an efficient sequencing method on duplex circuits because the individual channel duty cycles are only about 50%. Accordingly, the data rate for Plan A has to be about twice that of Plan B to handle the same quantity of information. An equivalent statement is that, for a given capacity, the data rate for single channel operation must be about twice that for dual channel, which is to be expected if the dual channel system is used efficiently.

6.3.2 Communication Coverage for Plan A

Consider now the problem of obtaining ground-to-air communication coverage over a large volume of air space when line of sight conditions prevail. Fig. 6.3 illustrates the geometry. Three transmitters are spaced on a great circle with a separation of d statute miles, and it is assumed that each transmitter provides communication coverage to all points above a plane tangent to the earth at the transmitter. There are zones of no coverage below the intersection of adjacent tangent planes, and zones of overlapping coverage above. It is clear that if adjacent transmitters are on the same frequency, interference will be encountered in the zone of overlapping coverage, which unfortunately begins at

the same altitude as full coverage.

The height limit h_1 at which overlapping coverage begins can be approximated by assuming that the line of sight radio path distance is the same as the great circle distance between ground points directly under the transmitter and receiver. This assumption is not nearly so rash in practice as might be imagined from Fig. 6.3 since the heights involved are actually a very small fraction of the earth radius. The line of sight radio path distance is given by[4]

$$D_L = \sqrt{2H_t} + \sqrt{2H_r} ,$$

where D_L is distance in statute miles, H_t and H_r are the heights of transmitting and receiving antennas respectively in feet, and a 4/3 correction is applied to the earth's radius to allow for average atmospheric refraction. Assuming a ground antenna height of 50 feet and a receiver height of h_1 feet,

$$D_L = 10 + \sqrt{2h_1} .$$

The geometry of Fig. 6.3 shows that

$$D_L \doteq \frac{d}{2} ,$$

where d is the transmitter separation, and the approximation is quite accurate when h_1 is small with respect to the earth's radius. It follows that

$$h_1 \doteq \frac{1}{2} \left(\frac{d}{2} - 10\right)^2 .$$

From similar considerations

$$h_2 \doteq \frac{1}{2}(d - 10)^2 ,$$

where h_2 is the height where overlapping coverage (or interference) begins for transmitters separated by 2d statute miles. Specifically, if h_2 is 100,000 feet, d is about 457 miles and the separation of non-interfering transmitters is 914 miles. It is obvious that much smaller separations are essential to obtain adequate coverage, and that interference will result at relatively low altitudes if any of these transmitters operate on the same frequency at the same time.

The interference problem can be eliminated for closely spaced transmitters on the same frequency by arranging matters so that no two emit at the same time. This requires control from some common point and is not difficult if the area covered is not too large. Suppose the area to be covered is represented by a square having sides of length d, where d is determined by the maximum height for interference free coverate h_2, that is,

$$d = 10 + \sqrt{2h_2} .$$

Within this square, enough transmitters will be placed to obtain adequate coverage. All will operate on the same frequency, airborne or ground, and interference will be avoided by controlling all transmissions from a common point. From the foregoing discussion, it is

clear that no boundaries of a similar square operating on the same frequency and with independent control should be closer than 2 d. From this beginning, a frequency chart for a large communications network can be built up as in Fig. 6.4. No more than 9 frequencies are required for complete coverage with such a system, since it can be extended in any direction. The dimensions of the individual squares are dependent on the height of interference-free coverage, and transmitters may be located anywhere within a square so long as they operate on the assigned frequency. Restrictions in the airborne case are similar. The line of sight range between two aircraft at 100,000 feet is 894 miles, or again about 2 d.

The foregoing coverage plan can provide two-way communication coverage in vertical sectors using a single channel per sector and is accordingly well suited for sequencing Plan A.

6.3.3 Communication Coverage for Plan B

Consider now a coverage plan suitable for providing communication in horizontal sectors or layers. Section 6.3.2 shows how interference-free coverage can be obtained by time-sharing a channel among several transmitters on the same frequency; consider now how frequency separation might be used to accomplish a similar result A con-

sideration of Fig. 6.3 shows that interference-free coverage can be obtained over a vertical band between the height limits h_1 and h_2 by alternating two transmitter frequencies, and it is not hard to see how four frequencies can be used to extend this coverage into a three dimensional shell. The idea is elaborated in Fig. 6.5 by means of a grid diagram similar to Fig. 6.4. Here coverage is obtained by alternating four frequencies f_1, f_2, f_3, and f_4 so that a minimum distance 2d separates all transmitters on the same frequency.

Unfortunately, the height boundaries are not as widely separated as we would like, as the following table shows:

Transmitter Spacing In Statute Miles (d)		Height Coverage in Feet h_1	h_2
436	(A layer)	21,600	100,000
218	(B layer)	4,900	21,600
109	(C layer)	990	4,900

It appears that three sets of four frequencies each are needed to provide complete ground-to-air coverage, superimposed as shown in the grid chart of Fig. 6.6. No transmitter switching is required; all transmitters may radiate simultaneously. However, a relatively large number of radio channels is required if conventional frequency spacing is used.

Voice coverage over very large areas on a single VHF channel has been achieved by the use of offset carrier methods[5,6,7,9]. These offer a kind of space and frequency diversity which may improve accuracy by 10 or 15 db[10], and a potential reliability improvement is inherent in the multiple transmitting facilities. Fig. 6.8 shows a possible frequency spacing diagram for offset carrier operation on a 50 kc channel. Four transmitting frequencies are spaced $\pm$ 7.5 and $\pm$ 15 kilocycles with respect to the nominal channel center frequency f_o. A VHF receiver tuned to this channel will pick up all four signals in its normal IF bandwidth, and the demodulated outputs will include a number of carrier beat frequencies, the lowest of which will be at 7.5 kc/s. A low-pass filter will suppress these beat notes in the audio output, while at the same time leaving a quite adequate audio bandwidth for voice modulation. If the same signal modulates all four carriers, coherent addition of the modulating signal can occur in the receiver output.

It appears, however, that the application of offset carrier methods to data transmission will require some caution. In general, the four signals f_1 to f_4 will arrive at an airborne receiver with different delays, and an unfortunate combination of delays and

signal strengths can result in a notch in the audio output. This is not normally serious in a wideband signal such as voice, but it could be disastrous for narrow band data signals. Frequency diversity within the modulation band, such as might be obtained by using different FSK sub-carrier signals on each carrier, would overcome this problem if the data rates were not too high. Special attention to such matters as detector linearity, frequency accuracy, and stability of the transmitters is also necessary. Since none of these problems is insurmountable, it is worthwhile to consider the radio channel requirements for a large two-way network using the offset carrier technique.

For ground-to-air transmission using offset carrier, it appears that three channels are sufficient to provide vertical coverage in three layers from 1000 to 100,000 feet. The modulation signals in each vertical layer may be unrelated, but a practical interference problem arises when it is necessary to provide independent modulation in adjacent horizontal areas within a layer. Fig. 6.7 illustrates the point.

A vertical cross section through the coverage pattern provided by a transmitter and frequency distribution like that shown in Fig. 6.5 would follow the curvature of the

earth as shown in Fig. 6.3. However, it can be represented as in Fig. 6.7 by a suitable distortion of the vertical scale. The areas of no coverage extending up to h_1 and the height of non-interfering coverage h_2 are shown for a single layer. Imagine now that the modulation applied to transmitter T_1 differs from the modulation applied to T_2, T_3, and T_4. In the cross-hatched area, coverage will be available from both modulation signals, and interference will result unless some way of separating them is adopted.

This separation in an offset carrier network can be achieved by the use of different sub-carrier frequencies for the two signals. The airborne data receiver can then be switched from one to the other at any point in the cross-hatched area of Fig. 6.7. It is desirable that such switching be ground-controlled so as to simplify co-ordination of sequence programming between the two modulation areas; this operation may well be simplified by the fact that an area exists in which coverage is available from both signals. There is no reason why the sub-carrier frequency at T_1 cannot be repeated at T_3 once the switch-over has been accomplished. One of the most attractive features of the method is that no RF channel changing is required; it appears that, so long as an aircraft stays within the height boundaries defined

by h_1 and h_2, it would be theoretically possible to fly around the world without retuning the airborne receiver. Further, world-wide ground-to-air coverage from 1,000 to 100,000 feet is theoretically possible on 3 channels.

An estimate of the air-to-ground channel requirements can be arrived at as follows. Suppose three air-to-ground channels are assigned to a block of air space with height 1,000 to 100,000 feet and horizontal cross section approximately equal to a square with sides equal to the line of sight distance for an aircraft operating at the top of the block. This assignment will provide equal air-to-ground and ground-to-air message capacity within the chosen block of air space. None of the three air-to-ground frequencies can be reassigned in neighboring blocks closer than twice the maximum line of sight distance if interference-free operation is to be obtained, so that the situation is analogous to that illustrated in Fig. 6.4. On this basis, it appears that a maximum of 27 air-to-ground channels will be needed. Unfortunately, the frequency stability of most airborne radio transmitters is not suitable for channel-saving schemes such as offset operation. It seems, therefore, that a total of 30 channels are needed for complete two-way coverage with the B sequencing plan

of Fig. 6.1; that is, about 3 times as many as for the A plan.

6.3.4 Ground Communications

In distributing data signals to widely separated ground transmitters, it is convenient to use common carrier voice networks. Unfortunately, these networks have not generally been designed with data signals in mind, and special care is necessary to achieve data rates commensurate with nominal line bandwidths. A typical problem arises because of the use of sideband methods in carrier telephone, where frequency and phase errors in introducing a locally generated carrier destroy the ability to preserve pulse shape at the receiver output. Sub-carrier modulation methods are normally used to overcome this problem.

However, at pulse rates above five or six hundred per second, frequency separation between the data signal and line sub-carrier is relatively small, resulting in considerable systematic pulse jitter at the output of the demodulator. Since the radio portion of the ground-to air path is not subject to the same limitations as the ground portion, it is probable that different modulation parameters will apply. Hence systematic jitter might be very small, but random jitter due to noise will be added. It is clearly desirable to minimize systematic

jitter before applying the signal to the radio transmitter, and to provide for a change in modulation parameters between the line and radio portions of the signal path. To put it another way, a clear-cut interface providing for synchronous regeneration between the line and radio components should be provided.

With regard to the influence of line characteristics on sequencing methods, the most significant factor appears to be the velocity of energy propagation along the line. In some cases this may be as little as 10,000 miles per second[14], so that delays up to 10 milliseconds may arise on a hundred mile line. This factor is unimportant in networks operating according to the A sequencing plan if each transmission can be treated as an independent event - that is, if it does not depend for synchronizing information on previous transmissions and has the effect of reducing the network capacity slightly. In the B sequencing plan, however, differential delays of this order in the signals emitted from two transmitters will result in something like HF multipath interference. In an offset carrier network, differential delays should be small with respect to the data pulse length, and delay equalization should be used. If synchronous regenerators are provided at the interfaces, delay insertion and adjustment will be relatively

simple, although the problems of delay equalization and calibration may be considerable.

6.4 Characteristics and Applications

6.4.1 Network Capacity

It has already been stated that sequencing Plan B as described in 6.2.2 is inherently capable of a higher data capacity than Plan A. This advantage can now be reviewed in terms of network capacity.

Plan A required 9 channels, and Plan B required 30. A 3 to 1 advantage accrues to the B Plan because of the greater number of channels, and an additional 2 to 1 advantage because of the methods of sequencing, so that the total capacity advantage of the B Plan is about 6 to 1. Since this is obtained at a cost ratio of about 3 to 1 in channels, all other factors being equal, it appears that the B Plan provides for fairly efficient use of its channel even though the total initial requirements are larger.

6.4.2 Radio Set Requirements

Most airborne VHF voice rad[illegible]n be adapted for A Plan service with virtually [illegible] modification, since the only [illegible]quirement is to transm[illegible]d receive alternately on the same channel. This is not true for the B Plan. Here the requirement is to receive on one of three

channels, and transmit on any one of 27 different channels, though again it is not necessary to transmit and receive simultaneously. To meet this requirement with most standard radios, retuning will be necessary before each transmission. When remotely controlled channel selection is available an automatic channel selector can be provided and the time required for channel changing can be accounted for in the sequence plan. If channel selection in the receiver and transmitter are independent, or if separate transmitters and receivers are available, the number of tuning operations will be greatly reduced.

6.4.3 Fault Conditions

In discussions about networks employing sequencing Plan A on a single radio channel, the so-called "stuck carrier" objection frequency comes up. The basis for concern on this point is the possibility that one of the network transmitters, through fault or human error, may be allowed to radiate continuously and thereby jam the network over most of the service area of the offending transmitter. Although such occurrences may be rare, they may represent a serious hazard potential and are difficult to protect against. Probably the safest procedure is to provide a separate clear channel to be used in an emergency. There is still the diffi-

culty that considerable time may elapse before all members of the network recognize the fault condition and switch to the alternate channel.

It is worth noting that the "stuck carrier" hazard is greatly reduced in the B sequencing plan. In the first place, the ground transmitters radiate continuously in normal operation, and the airborne transmitters operate on different frequencies, so that the "stuck carrier" problem does not arise in ground-to-air signalling. In the second place, multiple channels are normally provided for air-to-ground signalling, so that faults on one channel do not destroy all the normal air-to-ground capacity. In addition, since the ground-to-air path is not affected it is possible to switch all but the offending aircraft to a clear channel with a single general call message.

6.4.4 Extended Message Handling

There are two types of extended message to be considered: voice and data. Besides the obvious reasons, this distinction is convenient here because data messages can often be transmitted in a series of discontinuous bursts, whereas voice signals cannot be handled readily on a discontinuous channel.

Extended message transmission can be handled by either

serial or parallel methods. In the serial method, the sequencing program of the data link will be interrupted to provide time for transmission of the extended message; in the parallel method, the extended message will be transmitted in parallel with the regular program on a frequency multiplex basis and may share the synchronizing and address preamble with the normal message. This latter method is practical in a 3 kc channel if the basic data rate is not too high.

Since voice transmission on a discontinuous channel does not appear to be convenient, we may conclude that serial methods would be necessary in handling voice with sequencing Plan A. Digital headers and end of message codes should be provided on voice signals to assist in sequencing and control of the network. Extended data messages could be handled by either serial or parallel methods in sequencing Plan A, though it is likely that serial methods would be more convenient.

In sequencing Plan B, parallel transmission of extended messages is convenient on the ground-to-air path because of the continuous transmission. Since the airborne sets are maintained in continuous synchronism, it is convenient to transmit broadcast data in synchronism with the basic messages on a separate sub-carrier, so that any suitably equipped aircraft may monitor the broad-

cast messages at will by connecting a display device (such as a printer) to the appropriate data set output. In Plan B, extended messages on the air-to-ground path would be most conveniently handled on a separate radio channel. These will normally be voice, since it appears that most aircraft will not have the capability of composing extended data messages.

There appears to be some interest among airlines in a "company communications" capability. It goes without saying that if company communications is to share equipment and channels with the air traffic control service, it must do so without imposing any operational restrictions or significant degradation in performance. Consequently, ground-to-air company communications could not be handled as extended messages in the serial mode because of their effect on the control service timing. The parallel mode would have to be used, with either A or B requencing depending on whether the service required is predominately voice or binary data. On air-to-ground company messages, unless additional equipment is carried for composing general f .mat binary messages, voice will be used, probably on a separate channel reserved for the purpose. The B sequencing plan appears to offer more advantages for company communication purposes, because of greater capacity, and because of

flexibility in the assignment of air-to-ground channels and sequencing.

6.4.5 VOR Integration

Since VOR equipment is in common use and has a voice modulation capability which appears to be seldom used, it is worthwhile to consider the possibility of integrating ground-to-air data networks with VOR transmitters.

It is clear that, if VOR transmitters are used to convey data signals to aircraft, the air-to-ground responses must occur on a different frequency. Thus, sequencing Plan B can be considered, but a coverage plan similar to that of Fig. 6.4 would be applicable. All the aircraft in the service area of the VOR transmitter could be included in one sequencing program, but a separate program would be necessary for each VOR service area. Fairly frequent channel changes would be necessary and these would have to be coordinated and controlled from the ground.

The VOR signal includes an accurate reference tone which is recovered in the airborne receiver. An air-to-ground data system can be conceived in which this reference tone is used to provide bit synchronization over the VOR service area.

6.4.6 Pilot Initiated Messages

Pilot initiated messages may be considered under two general types: first, an emergency type in which little or no delay is tolerable; second, a non-emergency type in which delays of perhaps 2 minutes cause no serious inconvenience. They should not interfere with the control service and should not jam transmissions from other aircraft.

If the pilot-initiated message is not urgent, it can be delayed until normal sequential interrogation occurs and then transmitted as a normal extended message. Where such a delay is intolerable, some kind of break-in procedure to provide a clear air-to-ground channel is necessary.

One possible method might work as follows. Airborne sets can be designed so that no response occurs to a general call; this is not essential perhaps, but it might as well be done because simultaneous responses from all aircraft receiving a general call would be useless because of mutual interference. Now suppose that the airborne set can be made to respond to a general call when the pilot wishes to initiate a message. Two possible defects in such a scheme are (1) delays due to infrequent ground-to-air general calls and (2) the possibility of two or more air-

craft responding to the same call. Both of these difficulties can be minimized by transmitting a sufficient number of dummy general call messages, but the capacity of the network is reduced in proportion to the number of dummy messages.

In real emergencies, the pilot might simply pick up his microphone and start talking. The intruding transmission would immediately be apparent to ground control and although it might not be immediately intelligible due to interference from another aircraft making an automatic response, the duration of interference would probably not exceed one second, and the controller could manually intervene to interrupt the normal ground-to-air transmission temporarily. This kind of solution is not particularly elegant, but it does provide for emergency service with little delay and the loss of only one air-to-ground transmission. The emergency message will of course interrupt normal control service throughout its duration.

A third possibility for handling pilot-initiated messages is to provide a separate (guard) channel for them. The pilot might then simply tune his radio to the guard channel and in this way establish communication without reference to the control service. This kind of procedure might conceivably be used in a semi-urgent situation in

which the several seconds delay required to change radio channels is acceptable, and temporary disappearance of the automatic responses from the associated aircraft to control interrogations is tolerable. It is conceivable in fact that the pilot-initiated contact might be completed and the radio set retuned to the control channel with never a missed response.

6.4.7 Modulation

Tone shift signalling (FSK-AM) appears to offer a good compromise between the many conflicting requirements. The main reason for this choice is the requirement for operating the data communication system over radio and line facilities primarily intended for voice service. These voice facilities are already in service in air traffic control communications, and it seems unlikely that a completely independent service for data will be practical in the near term. Extensive modification of existing facilities is similarly undesirable. Fortunately, tone shift signalling can be integrated with voice networks with little difficulty. It is simple, easily serviced, and presents no serious frequency stability problems in AM radio equipment. In SSB equipment, the picture is not quite so favorable due to doppler shift and frequency inaccuracies, but there is some compensation from the high power efficiency which can be real-

ized. In any case, there are few simple pulse modulation schemes suitable for use on SSB radios, and tone shift modulation appears to be one of the simplest.

From the results presented in Section 5 of this report it may be concluded that efficient FSK signalling can be accomplished in a bandwidth equal to about one half times the reciprocal of the pulse length, and with total shifts of about half as much, provided that 8 or more tone cycles per pulse are transmitted. Expressing these results in algebraic form, we may write for the lower of the two signalling frequencies

$$f_1 = \frac{8}{T}$$

and after Kotel'nikov[12] for the shift

$$\Delta f = \frac{.7}{T},$$

so that the upper signalling frequency is

$$f_2 = f_1 + \Delta f = \frac{8.7}{T},$$

where T is the pulse length. The channel center frequency is given by

$$f_o = \frac{f_1 + f_2}{2} = \frac{16.7}{2T},$$

and the signalling bandwidth is

$$f_B = \frac{1.5}{T}.$$

The upper edge of the signalling band is given by the

sum of the center frequency f_o and half the signalling bandwidth f_B. Equating this quantity to the upper limit value of 3000 cps and solving for T, it turns out that the pulse length for optimum performance is approximately

$$T = 3.04 \text{ milliseconds},$$

which corresponds to a rate of about 330 pulses per second. The lower limit of the data signalling band occurs at $f_o - \frac{f_B}{2}$ or about 2500 cps. The lower part of the voice channel is adequate for voice or other low speed data.

The result obtained in the foregoing paragraph may seem surprisingly low in view of the available bandwidth. One of the critical factors is the upper frequency limit of the signalling band; the other is the requirement for no less than 8 sub-carrier cycles per bit. No experimental evidence has been obtained to indicate what degradation in performance might result for lower sub-carrier frequencies in FSK-AM, although an indication can be deduced from the influence of timing inaccuracies on error rates. See Fig. 5.16. It appears that timing inaccuracies above about 20% result in a rapid rise of the bit error rate. Systematic jitter of this magnitude will occur in a simple FSK-AM system at 5 of 6 cycles per bit where the keyer frequency is not synchronously related to the bit rate. Such jitter is not particularly critical when the

average signal-to-noise ratio is large, as on a typical telephone circuit, but it will be very damaging on a noisy radio circuit.

Long range data communication in the HF band requires special consideration due to pulse smearing and selective fading introduced by multipath conditions. The signal arrives at the receiver by several different routes of varying lengths and attenuations, so that the receiver input sees a composite signal in which the relative phases and attenuations of the individual components may vary at quite high rates. Path length variations smear the signal transitions. This problem is normally overcome by using data pulse lengths which are long with respect to the differential delays on the HF path. Law[10] and others have reported that differential delays on an HF path seldom exceed 5 milliseconds, and typical values are closer to 2 or 3 milliseconds. As a consequence, data pulse lengths less than about 10 milliseconds are seldom used on an HF path.

Selective fading arises from the cancellation of signal components, the degree of cancellation depending on the relative attenuation of different paths and the frequency at which cancellation occurs on the differential path lengths. These fades are generally uncorrelated at frequencies separated by as little as 500 cps, so that

frequency diversity is useful in minimizing their effects.

A simple and effective modulation method for long range data communication in the HF band is provided by two FSK-AM sub-carriers, separated by several hundred cycles, each modulated by the same data signal. This method provides the necessary frequency diversity, and the use of adequate pulse lengths protects against smearing.

6.4.8 Coding

Suppose that a set of several independent variables is to be encoded in a binary sequence and that the resolution and range of values for each variable is specified. The specifications will indicate the minimum number of binary symbols required to represent each variable according to the inequality

$$n-1 < \log_2 K_i \leq n \, ,$$

where K_i represents the number of discrete values required to define the i^{th} variable and n is the smallest integer which satisfies the inequality. In general the value of n will be different for each variable and the 2^n combinations which can be formed from n binary digits will exceed the number of discrete values to be represented. A typical example arises when the variable to be encoded is a decimal number. In this case, we have

$$n-1 < \log_2 10 \leq n \, ,$$

and the smallest integer which satisfies the inequality is n = 4. A binary sequence having a resolution of $2^4 = 16$ values is needed, so that there will be six excess code symbols. If the problem called for 17 discrete values, there would be 15 excess code symbols, and though the situation would not normally provide such a poor fit we cannot expect all communication requirements to resolve themselves neatly into powers of two.

It should not be inferred here that excess capacity is of no value; on the contrary, it can often be put to use in some form of error checking. For instance, we may select a five-bit code to represent decimal numbers, choosing only those members which contain exactly two pulses. Out of the 32 code combinations, exactly ten will satisfy the two-pulse requirement, and if any of the remaining 22 are received, they can be recognized as errors because the two-pulse requirement is not satisfied. The two out of five code provides an example of putting excess capacity to good use. There are other examples, such as binary-coded decimal, where the excess capacity is difficult to utilize, and it appears that codes designed for the "closest fit" to the actual requirement (such as binary-coded decimal) fall into this group.

A coding method in which all variables can be represented

by a single set of equal length codes is preferred in a majority of automatic data systems because it confers the following advantages:

- Machine timing and synchronization is simpler.
- The coding problem reduces to a single choice.
- Controlled redundance can be used for error correction.
- Simple error detection methods can be devised.
- Performance can be predicted readily because the error probabilities for each code member are equal.
- Code translation is generally easier between systems employing equal length codes. Inter-system coupling problems are minimized.
- Message format changes can be made easily.
- Arithmetic operations can be performed more easily.

These advantages can be roughly summarized in two categories--compatibility and simplicity--both of which are of prime importance in the air traffic control data link. Since a close fit between the message requirements on the

one hand and binary coding on the other is difficult to obtain, it appears that there is little to lose and much to gain by adopting a fixed length code.

6.4.9 Error Correction

Communication devices in a difficult environment are often operated with a considerable safety margin[17] to protect against temporary or unforeseen adversities. The term "error free" has been applied to such systems, but it is obvious that a literal interpretation of such terminology is impossible. A system which never makes an error cannot be made, and we would never know if we had one if we watched it assiduously all our lives. We must confine our attentions to practical systems having finite error rates, however small. In the case of an air traffic control data link, defining an acceptable error rate may be unpalatable because the objective is to assure safety, but it is nevertheless inescapable if the requirements are to be analyzed at all.

Having defined an acceptable error rate, we may consider the communications environment to determine how it is to be obtained. System margins can be increased by lowering the pulse rate, all other factors being held constant, when the message capacity requirements permit. If the objective cannot be attained by a direct approach within the limitations of available signalling power,

efficient design, and practical economics, it is possible that error correction methods will provide a satisfactory solution.

Error correction may be obtained by re-transmission of erroneous data, or by real-time self-correcting methods[15]. Re-transmission methods tend to be relatively simple and fit in well with the two-way communications requirement. An objection is that the information rate slows down under difficult conditions due to the increased number of re-transmissions, reducing the system capacity. This factor can be balanced against the situation using self-error-correcting codes of the type described by Hamming[1], where the information rate is determined at all times by reference to the most difficult conditions likely to be encountered, and there is no provision for automatically taking advantage of favorable conditions by increasing it. For an idea of magnitudes involved in a specific case, consider a single-error correcting code and assume that five binary symbols represent the information content of the code. In the case of correction by re-transmission, let the code include enough redundance to permit the detection of single errors; this capability can be provided by adding a single parity bit to each code member. In the case of a single error self-correcting code, Hamming has shown that four redundant bits are required for each

five information bits. If the bit rate is the same for each code, the information rate of the re-transmission method may approach 1.5 times that of the self-correcting method under favorable conditions. Under adverse conditions the comparison is not so easily made because the length of repeated messages must be considered[16]. The best rate is achieved when only those characters which are incorrectly received are re-transmitted; this kind of operation requires full duplex service and variable message length.

Correction by re-transmission can be no more effective than the error detection capability of the code. It has been shown in Section 5 that perfect correction based on a single parity check over 5 bits will yield about the same reduction in error rates as a 3 db increase in signal power. The advantage will increase with more redundant error detecting codes, of which there are great numbers.

A method of self-error correction due to Wagner[2] exacts no penalty in information rate, but does increase the receiver complexity. Instead of providing complete correction, this method employs analogue information derived from the received signal to maximize the probability that correct decisions are produced at the receiver. Null zone detection methods appear to be a degenerate form of

Wagner's basic technique, and results presented earlier in this report indicate that their effectiveness is comparatively low. Improvements obtained with Wagner decoding are apparently due to efficient use of amplitude information in the received signal, which is discarded in conventional FSK receivers and used only in a limited way in null zone receivers.

Some recent reports[18] indicate that error probability on land lines is about 1 part in 10^5 with most errors occurring in clusters due to bursts of interference. Since many of these error clusters are quite long, a self-correcting code capable of dealing satisfactorily with them would have to be correspondingly long and very complex. It seems to be generally agreed that correction by re-transmission is the simplest way to reduce line error rates.

6.5 Influence of Message Format

6.5.1 Typical Format

The behavior of communication networks depends on the detailed structure of the transmissions themselves as well as the sequencing plan and characteristics of the equipment. Fig. 6.2 represents some of the factors of a general message in a time sequence diagram. The first step, getting the transmitter on the air, begins with a switching command which may originate at a push-to-talk

button or an automatic control circuit and ends when the transmitter has settled into a steady state and is ready to accept modulation. The time interval involved here may be considerable. In most cases, the transmitter will maintain tube heaters at operating temperature whether it is radiating or not; if this is not done, delays in the order of tens of seconds will be encountered. Generally, the delay interval includes the sum of the operating time of the transmit-receive switching relay and the build-up time of the high voltage power supply and has magnitudes from tens to hundreds of milliseconds. When an antenna change-over relay is included, as in common antenna working, its switching time may generally be ignored, because it is less than the high voltage switching and build-up delay and can be initiated at the same time by a parallel command.

When the transmitter switching transient decays to a negligible value, modulation can be applied. For reasons of convenience, modulation might be applied at the same time as the transmitter switching command so that some qualification based on sideband power output might be applicable here, but for purposes of discussion it is more convenient to adopt the simplification shown in Fig. 6.2.

It is not possible in general to begin transmitting the message immediately, particularly when data transmission

is involved. A synchronizing preamble followed by a message start may be necessary, and for both voice and data modes an address code has to be transmitted. Following this, the message proper may begin. Fixed or general format messages, or a combination of both may be sent, and though Fig. 6.2 shows both types in time sequence it is also possible that they might be sent in parallel. The end of the message proper may be marked off by an end of message code, and finally the modulation terminates and the transmitter switches off. Again, a transmitter switching transient occurs. The duration here is likely to be short--a few tens of milliseconds--and is mainly due to the release of the control relays and the decay of the transmitter high voltage supply. The "off" transient may be considered as ended when a nearby receiver on the same frequency reaches acceptable sensitivity.

A consideration of the foregoing time factors shows that a sizeable part of each transmission has little to do with the actual transfer of new information, viz., the switching transients and the synchronizing preamble. The identity code itself, or part of it, may be unnecessary in certain air-to-ground transmissions, as may be noted by considering current voice practice. The full identity may be used in establishing initial contact,

and an abbreviated form used in the subsequent acknowledgement.

6.5.2 The Switching Transient

When sequencing Plan A is used on a single radio channel, there does not seem to be any way to eliminate time delays due to transmitter switching transients, although it is certain possible to arrange matters so that the switch-on transient of one transmitter coincides in part with the switch-off transient of another. Since the switch-on transient is probably longer, it should be safe to assume that the total lost time per channel is proportional to the number of transmissions and the average switch-on transient time likely to be encountered. Assuming 125 message exchanges per minute and 30 milliseconds switching time for each transmission, the total lost time per minute is 60 x 125 milliseconds or about 7.5 seconds--somewhat over 12%. Many existing airborne transmitters will probably require a longer switching time, so that the above estimate may be low. However, it is not necessary to require that the switching delay allowance be equal for all sets; instead, it should be adjusted to match the individual sets by providing for delay increments in intervals of one character. In this way, lost time can be minimized without excluding slow-switching transmitters, and without upsetting pre-established

character synchronization in the network.

Sequencing Plan B, although it does not require any switching of the ground transmitters, does not materially reduce lost time because the airborne transmitters must still be switched.

6.5.3 The Synchronizing Preamble

Consider now the case of the synchronizing preamble. Systems based on the message format of Fig. 6.2 will be very flexible indeed, since every transmission is an independent event in the sense that no receiver requires any information from a prior message to help it interpret a new one. This is an essential characteristic when a receiver can detect only those messages which are specifically directed to it, but such is not always the case in an air traffic control communication network. The airborne receiver will overhear many messages directed to other aircraft, although it will not act on them. It may be possible to use them to maintain synchronization, so that the preamble of Fig. 6.2 can be eliminated except perhaps in certain special cases; if so, a significant increase in network capacity will result.

In single channel coverage such as might be used with sequencing Plan A, each airborne receiver may pick up signals from several widely separated transmitters in

sequence. Even if these signals were emitted in precise synchronization, they would arrive at each airborne receiver at different times due to the various radio path lengths involved, so that a time reference derived from one would not be optimum for another. For a maximum path difference of 200 miles, the timing error would not exceed one millisecond. The significance of timing error in a binary data receiver is directly related to pulse length; roughly speaking, timing errors of $\pm$ 10% may be considered negligible. See Fig. 5.16. It follows that satisfactory timing accuracy could be maintained in a type A system without the use of sunchronizing preambles if the pulse length were about 10 milliseconds. Further, if a common time reference could be established throughout the network, it would not really be essential for the ground receiver to readjust its synchronization on each air-to-ground message, so that the radio site might use a common time base for synchronous signal regeneration at both its transmitting and receiving interfaces.

Elimination of the sync preamble on the ground-to-air signal will require inherently good time base accuracy if synchronization must be maintained for considerable periods without correction. If the time base oscillator accuracy is taken as a 1 part in 10^6, a synchronizing

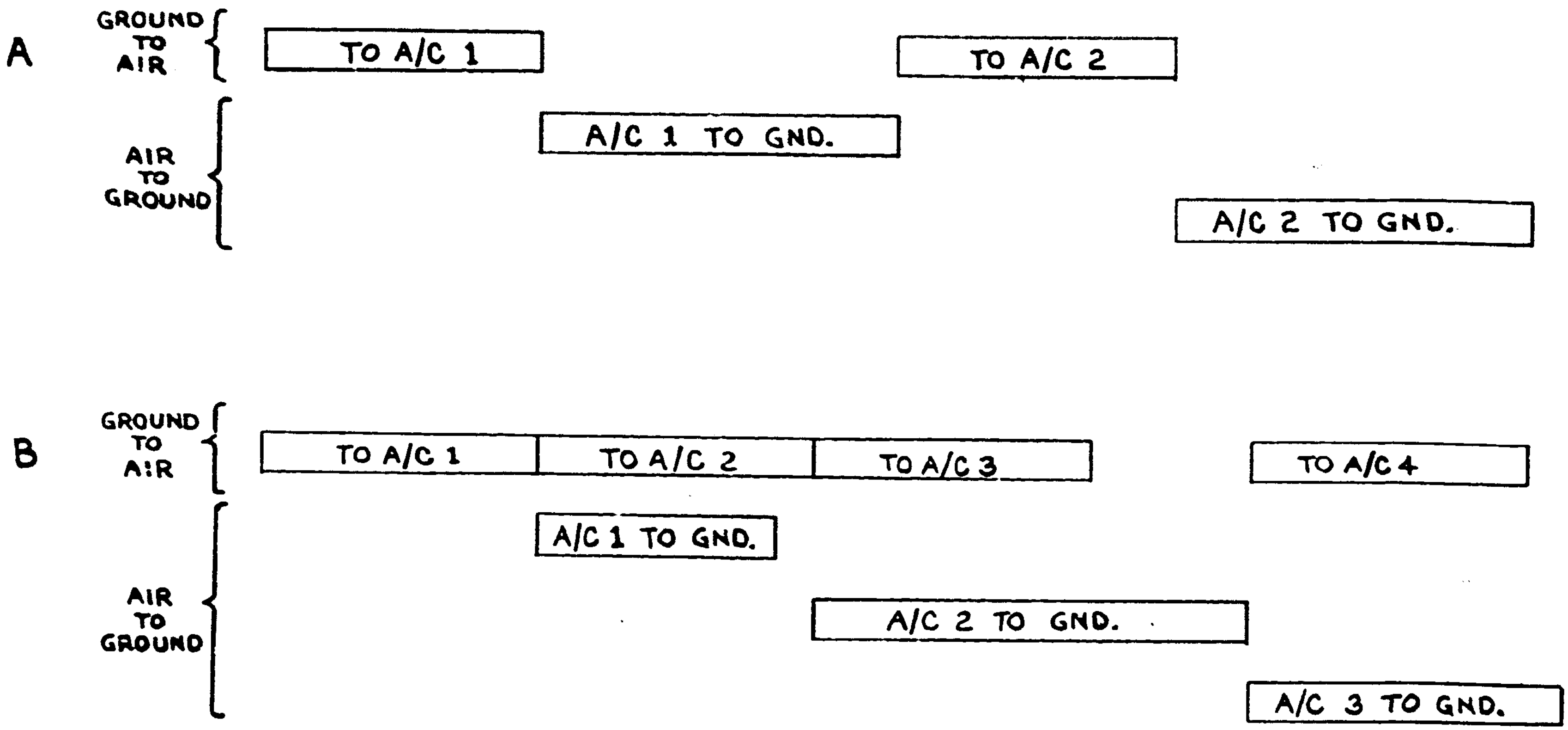
error of 1/10 bit may develop if no correction is provided over 10^5 bits. The maximum time which may elapse between corrections is accordingly dependent on the bit rate and will be about 100 seconds at 1000 bits per second, and about 33 minutes at 50 bits per second.

Some care in programming the ground-to-air transmissions is necessary to ensure that no part of the network service area is ignored for more than about 10^5 bits. Dummy transmissions, such as confidence checks, should be transmitted if no data is available for transfer.

With a continuous time reference throughout the network service area, it is possible to eliminate the start and end of message codes of Fig. 6.2 if all messages are strictly confined to a single fixed format or block length. The above methods have been applied in the data test sets described in Section 1. Although these sets are basically intended for use in a network using sequencing Plan B, it appears that the methods are general enough to be applicable in networks operating according to Plan A.

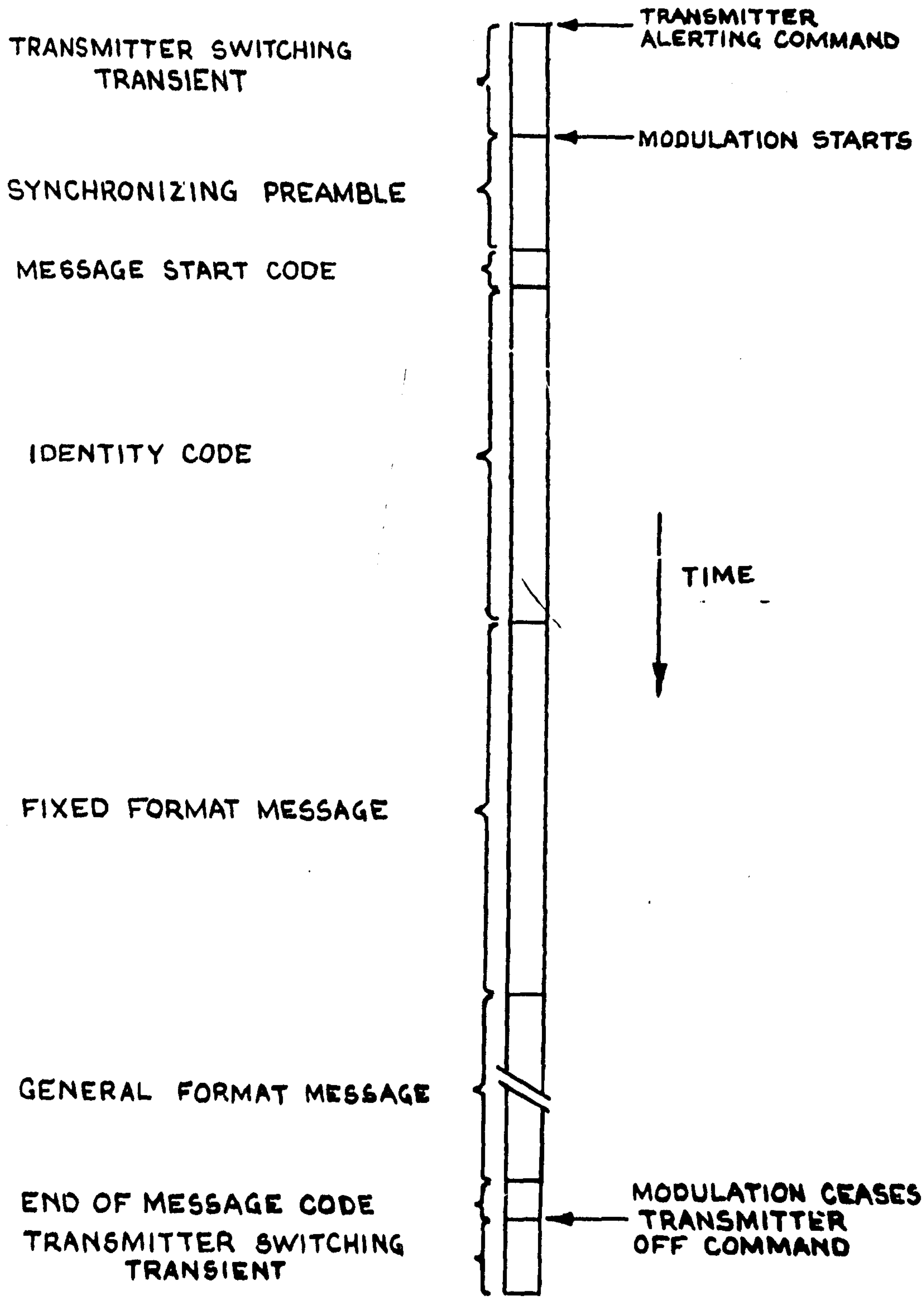
It does not appear that elimination of the synchronizing preamble permits any significant simplification of the data sets or destroys the capability of working on single isolated transmissions where synchronizing preambles must be employed. The reason for this is that the ability to

synchronize must be built into the set, but the actual details of message format by which synchronization is accomplished do not materially affect the design. The advantage to be gained is in increased network capacity due to the elimination of repetitions or non-functional synchronizing information, and it can be obtained by choosing suitable pulse lengths, providing stable timing generators, and equalizing line delays by simple digital methods.



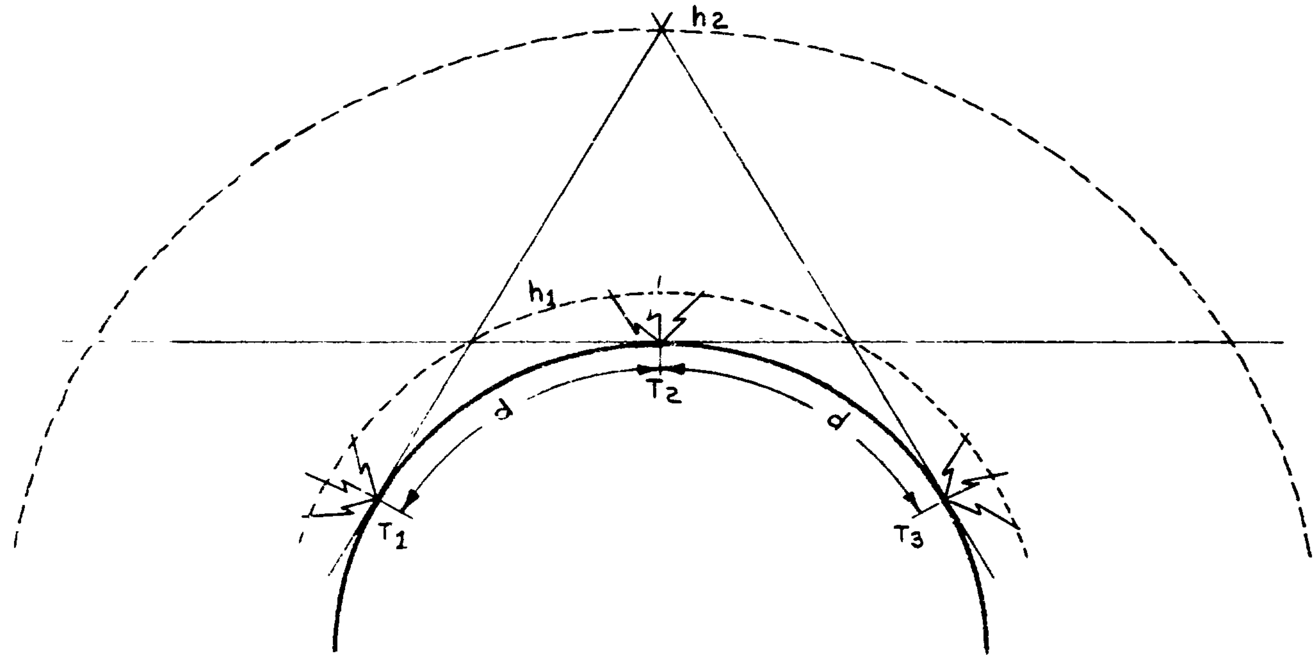
SEQUENCING DIAGRAM

FIG. 6.1



TYPICAL MESSAGE FORMAT

FIG. 6.2



COMMUNICATION COVERAGE DIAGRAM

FIG. 6.3

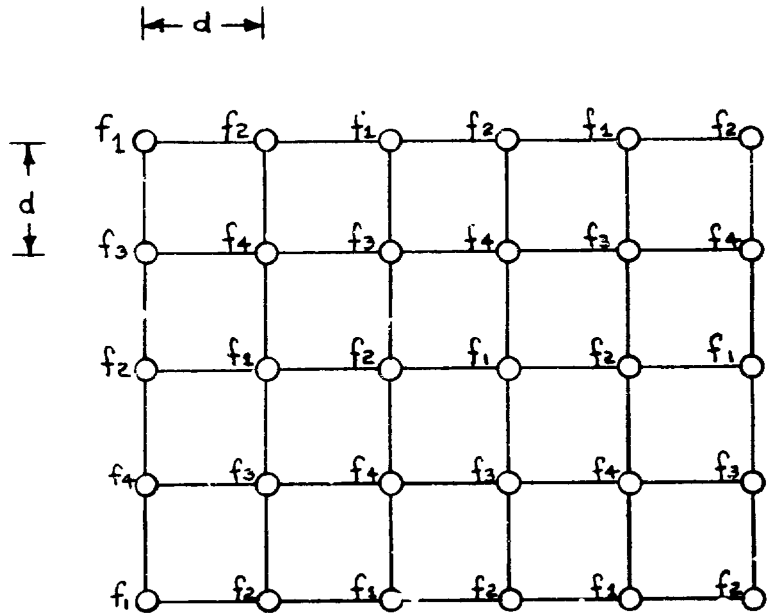
		d	2d	3d
	f_1	f_2	f_3	f_1
d	f_4	f_5	f_6	f_4
2d	f_7	f_8	f_9	f_7
3d	f_1	f_2	f_3	f_1

$$d\ (\text{MILES}) = 10 + \sqrt{2h_2}\ (\text{FEET})$$

WHERE h_2 = MAX. HEIGHT OF INTERFERENCE FREE COVERAGE

INTERFERENCE — FREE FREQUENCY ALLOCATION

FIG. 6.4



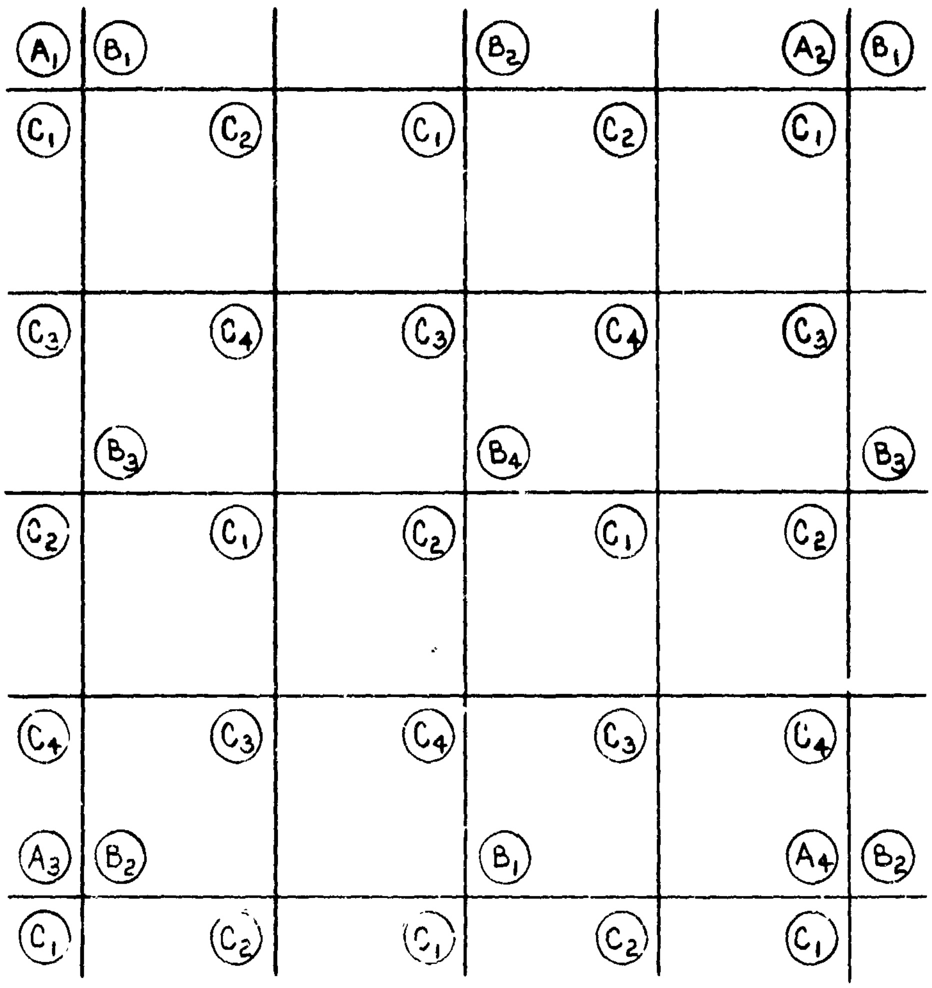
$$d\,(\text{MILES}) = 10 + \sqrt{2h_2} \;\; \text{FEET}$$

WHERE $h_2 =$ MAX. HEIGHT FOR INTERFERENCE FREE COVERAGE.

FREQUENCY DISTRIBUTION
FOR LAYER COVERAGE

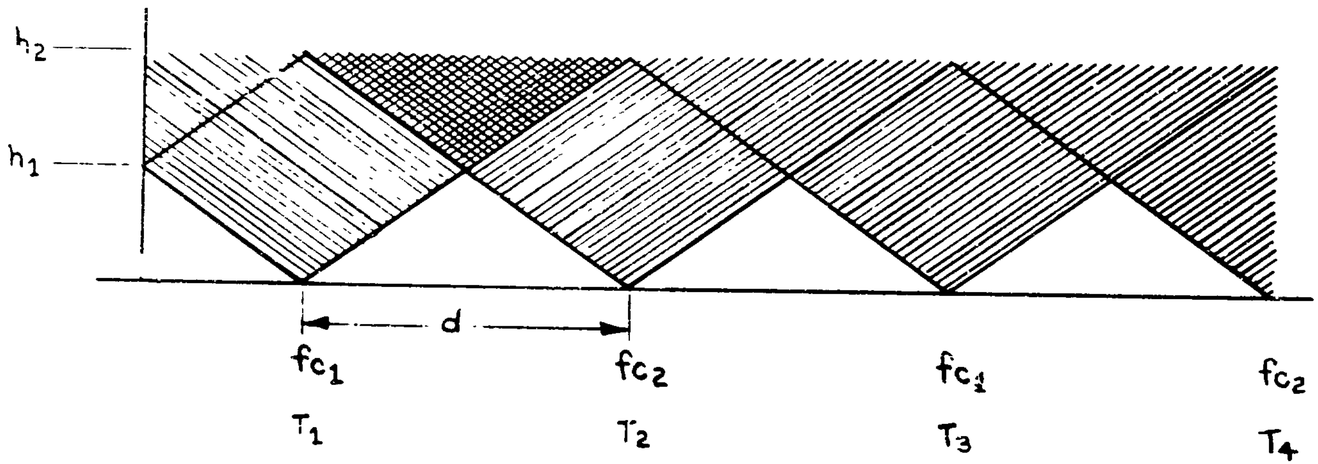
FIG. 6.5

A — UPPER LAYER
B — MIDDLE LAYER
C — LOWER LAYER



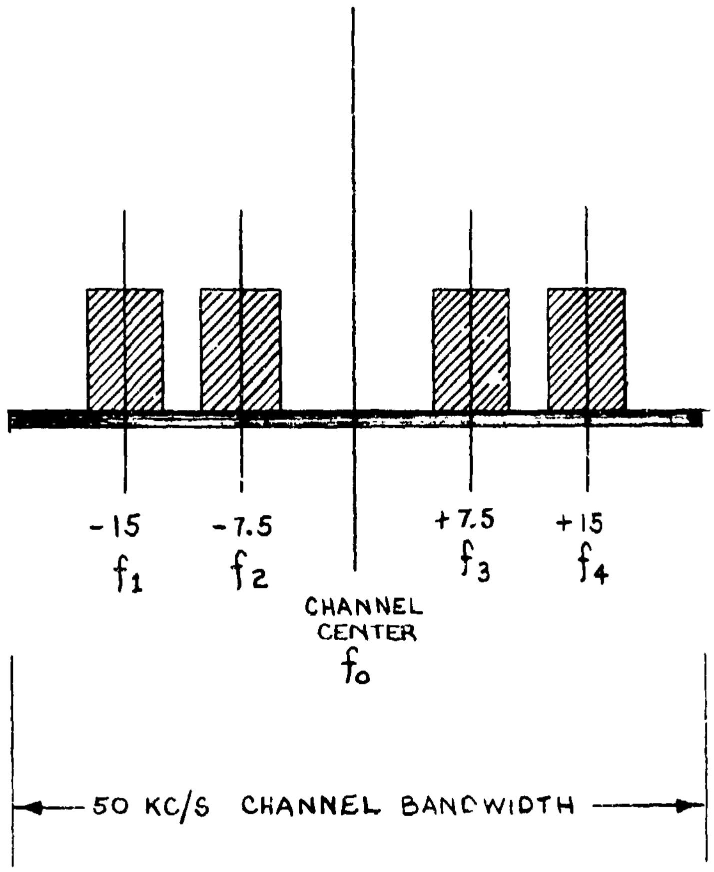
FREQUENCY DISTRIBUTION FOR
THREE LAYER COVERAGE

FIG. 6.6



INTERFERENCE BETWEEN ZONES OF INDEPENDENT MODULATION

FIG. 6.7



OFFSET CARRIER SPACING

FIG. 6.8

SECTION 7.0 CONCLUSION AND RECOMMENDATIONS

7.1 Sub-carrier Signalling Characteristics

7.1.1 FSK-AM

The results of Section 5 show that the performance of simple FSK-AM modems using phase shift discriminator circuits can be made to approach the theoretical predictions of Montgomery very closely under laboratory conditions. Predicted performance can be obtained by (1) choosing appropriate pre-detection and post-detection filtering, (2) employing synchronously timed sampling of the demodulator output. It is shown that Goldman's filter bandwidth criterion for maximizing SNR of pulsed signals leads to near optimum results when interpreted in terms of the 3 db bandwidth of relatively simple (2 branch) filters and that Kotel'-nikov's criterion for frequency shift similarly yields near-optimum results. There is a possibility that minimum shift (and bandwidth) may not be compatible with dual filter or Travis type discriminators because of inter-symbol interference.

The upper frequency limit of the signalling band provides a severe limitation on the maximum speed for FSK-AM signalling, primarily because typical discriminator circuits require several cycles of sub-carrier in each pulsed sinusoid to determine its frequency. It has been

shown that practical circuits can operate with as little as 8 cycles of sub-carrier per pulse without appreciable performance loss. The use of fewer sub-carrier cycles per bit will lead to an increase in systematic jitter at the converter output for FSK-AM signalling modems using non-coherent keying. Some experimental evidence on the influence of timing inaccuracy on bit error rates suggests that 10% of systematic jitter will degrade error rate appreciably. The jitter magnitude of a keyed two-tone signal has not been determined but it is reasonable to assume that it would not be less than a half period of the lower signalling frequency. Accordingly, it appears that the signalling parameters for FSK-AM should be chosen to provide not less than 5 sub-carrier cycles per bit to avoid significant performance degradation.

With this limitation in mind, it is useful to re-examine the maximum signalling rate for FSK-AM with an upper frequency limit of 3kc/s. Following the method of Section 6.4.7, we find that a signalling rate of 150 bits per second can be supported without appreciable performance degradation. The lower limit of the data signalling band falls at about 2250 cps, so that a nominal 3 kc/s voice channel can provide intelligible voice plus 500 bits per second FSK-AM data. Higher data rates can

be obtained only by sacrificing data performance through the use of fewer cycles per bit, by raising the upper frequency limit of the signalling band, or by increasing modem complexity to provide for coherent keying and more sophisticated demodulators.

7.1.2 PSK-AM

It is interesting to note that PSK-AM can be used to increase data signalling rates in a voice channel because good performance can be obtained with only one sub-carrier cycle per bit, using coherent keying. The signalling function in this case may be regarded as a cosine squared pulse having a base width T, so that a series of like pulses (such as all ones, or all zeros) will appear as an audio tone at the sub-carrier frequency $\frac{1}{T}$ cps. Nearly all the energy in this pulse will be contained in a bandwidth of $\frac{2}{T}$ cps, corresponding to the separation between the first zeros in the pulse frequency distribution, and will be symmetrically disposed with respect to the sub-carrier frequency $\frac{1}{T}$. Thus, the maximum signalling speed for a channel having an upper frequency limit of 3 kc/s is 1500 bits per second, and significant signal components will exist down to zero frequency. This fact will result in distortion of the signal in most voice channels because of restricted low frequency response; however, the at-

tendant performance degradation should not be serious because of the relatively small amount of energy in the extremes of the signalling band.

7.1.3 Bandwidth Considerations

The frequency spectrum of a PSK signal or a pulsed sinusoid keyed at a rate of $\frac{1}{T}$ bits per second is a $\frac{\sin x}{x}$ distribution with most of the energy contained in a bandwidth or $\frac{2}{T}$ cps. The minimum shift FSK signal, considered as the superposition of two pulsed sinusoids, occupies additional bandwidth due to the frequency displacement of its two components of about $\frac{.7}{T}$, making the total bandwidth about $\frac{2.7}{T}$ cps. However, the results of Section 2 show that the performance of an FSK demodulator using a nominal bandwidth of $\frac{1.5}{T}$ cps is not more than 2 db worse than that of a PSK modem using a matched filter. It seems likely that these results would not be significantly altered if the bandwidth of both signals were restricted to $\frac{2}{T}$ before transmission. The conclusion is that the bandwidth requirements of PSK-AM and FSK-AM are not significantly different from the detection point of view. Appreciable degradation in PSK performance can be expected if the transmission bandwidth for PSK is restricted to less than $\frac{2}{T}$ cycles per second because the phase transitions will be slowed or "smeared" in the filter output. In addition, the use of a derived refer-

ence, with its inevitable noise content, will further degrade PSK performance so that very little of the theoretical advantages over FSK will be realized in practice.

A phase ambiguity of π radians is inherent in the derived reference of a $\pi/2$ PSK receiver. Various methods may be used to overcome it, including transmission of additional synchronizing tones, superposition of synchronizing information on the FSK sub-carrier, detection of inverted code patterns, or polyphase modulation. All involve some sacrifice in either signalling power, time, or bandwidth, although it is frequently possible to minimize losses by using message synchronizing preambles to resolve sub-carrier reference ambiguity as well as to provide bit and character synchronization.

7.1.4 Error Rate

The results of Section 5.1 show that the bit error rate for FSK can, in practice, be accurately defined in terms of the SNR presented to the limiter of an optimized converter, and that it is independent of pulse rate. It does not follow that error rate is independent of pulse rate for a fixed power transmission, for the reason that signalling bandwidth will increase as the pulse length decreases, and consequently the

SNR presented to the limiters will decrease. The results of Section 5.1 clearly substantiate Montgomery's equation (Sec. 5.1, eq.1) relating probability of error and SNR. It follows that the relation derived from it between probability of error and pulse length in fixed power systems (Sec. 5.1, eq. 6) is also valid. Equation (8) of Section 5.1 summarizes the relationship between pulse length and performance in fixed power FSK systems, conveniently and simply. As far as error rate is concerned, the performance of FSK systems using different pulse lengths over like paths will be equivalent if the power transmitted is inversely proportional to pulse length.

Fig. 5.7 shows a typical relation between bit error rate and pulse length for an FSK converter. Inspection of this curve shows that the error rate is relatively constant for long pulses and rises steeply as the pulse length shortens. The optimum pulse length is defined by the point of tangency between the curve of actual performance and the curve of optimum performance as defined by equation (6) of Section 5.1. The curves show that it is better to err in the direction of longer pulses if the converter cannot be worked at the optimum point.

7.2 Communication Network Considerations

7.2.1 Design Parameters

In a communication problem involving both radio and wire services, there are many interdependent design considerations. Section 6 presents some of the implications inherent in the choice of a sequencing plan. The results of Section 5 are mainly applicable to the radio portions of the network. Other factors of importance to network performance and characteristics are the topology of the wire network, voice and data integration methods, compatibility with voice procedures, reliability requirements, radio frequency assignments, and radio frequency coverage. The design of the communication system must take account of interdependencies between these factors, and the detail design of the terminal equipment cannot be properly completed until they are resolved.

7.2.2 Network Topology

An example of the far-reaching implications inherent in basic system parameters is provided by the interdependence between the wire network topology and the system message formats. Consider a basic wire network consisting of radial feeders connecting a communication center to an array of remote transmitting sites. In this case, the communications center might

f rmulate messages and route them by performing a line selection operation. However, a more general (and in all probability a more economical) wire network might consist of a limited number of branched radials, and in this case message routing cannot be performed entirely at the center. It is necessary to provide in the message format a site selection preamble which will allow the message to find its way to the proper transmitter. It should be noted that some form of preamble is necessary in any case when the system is based on time-sharing a single channel between a group of transmitters, because the appropriate transmitter must be turned on to relay the message, preferably in real time. In a radial network, this might be done simply by voice-operated relay methods, so that the preamble might be analogous to a preliminary throat-clearing preceding the main announcement; in a branched network, a more elaborate preamble including a site code would be necessary, and it would have to be used whether a voice or data transmission were comtemplated.

The transmission and routing of signals from a communication center over a network is one side of the problem; it is also necessary to accept and route messages from any part of the network to the center. If the system operated invariably on a call and response basis, the

path of the out-going message could be maintained for the response; in general, however, it is necessary to originate certain messages (such as emergency voice) at random, in which case no pre-determined path will be available. It appears that a radial feeder network without branches will minimize the problems of routing randomly-originated incoming messages.

7.2.3 Interfaces

In Section 6.3.4, a case for the use of synchronous regeneration at the interface between the wire and radio portions of the network was briefly presented and justified on the grounds of efficiency for data communications. Signal regeneration at intermediate points in a communication network has been employed to advantage for many years, notably in telegraph signalling, but it should be noted that the insertion of digital signal regenerators in a network intended to provide both voice and data service indiscriminately presents a design problem. The presence of digital regenerators will effectively block the passage of uncoded voice signals, so that each regenerator must be by-passed for voice signals. It is clear that some form of automatic switching control is necessary to select the voice or datapath through the interface. Since the digital and voice signals fall in the same

frequency range, and since the voice signals may in general appear at random times, it is also clear that the use of automatic switching will require special attention, perhaps to the extent of modifying data message formats.

7.2.4 Random Access

Section 6.4.6 raises the possibility that the provision of random-access time-shared voice service on a centrally controlled data network, without the mutilation of occasional data messages or the introduction of undesirable delays, may entail a considerable sacrifice in network capacity. A further operational difficulty arises when the voice and data signals occupy a common frequency band, viz., the elimination of data signalling interference from the voice service. This particular problem can be solved in the airborne terminal by suppressing the voice output for a short interval after the receiver unsquelches. During this interval, the incoming signal may be analyzed to determine whether it is voice or data. If it is data, suppression of the voice output may be continued for the duration of the message.

In the case of data signals, or voice signals originating from the control center, the network behavior is orderly and predictable because of the close control

exercised by the communications center. Thus a communication path through the network can be selected and held by the center to serve both the outgoing message and the response to it. It is necessary to monitor only that line on which the response is expected. The situation in respect to randomly originated air-to-ground messages is considerably more chaotic. It is impossible to establish in advance a network path for them and it is also impossible to anticipate which line must be monitored to receive them. Hence, all inputs to the center must be monitored continuously for randomly originated messages. If the network is a large one, with a number of radio receivers tuned to operate on the common channel, the problem at the center can be seen to consist of two parts: first, several ground receivers may pick up an outgoing signal and attempt to present it to the central monitor; second, a randomly originated airborne transmission may be picked up and presented to the central monitor via several receivers.

Suppose that the control monitor is arranged so that all outputs are suppressed during a ground-to-air transmission and that no ground-to-air transmissions are permitted when any one of the ground receivers is not squelched. Under these conditions, a randomly

originated airborne transmission may seize control of the network at any time when the communication center is off the air, and a clear channel will be available so long as there are no interfering air-to-ground transmissions. The situation is very similar to current voice practice and tends to minimize the probability of interference between ground-based and airborne transmissions.

If a randomly originated air-to-ground signal is picked up by several receivers and presented to the control monitor, a situation akin to space diversity based on signal addition prevails, and cancellation of certain frequency components may occur. Section 6.3.3 discusses a similar problem in connection with off-set carrier methods. It is concluded that the combined signal will very probably provide satisfactory voice service, but that its suitability for data service is questionable.

7.2.5 Reliability

It is sometimes assumed that temporary failure of all or part of an automatic data communication network for air traffic control is tolerable if the system provides the capability of reverting to voice service on the same network. This assumption is not warranted in every case, as the following example will show.

Suppose that the data communication network is designed to serve a relatively large control area on a single radio channel basis. This assumption appears reasonable enough, since a considerable increase in channel capacity ought to result from the increased efficiency of automatic high speed communication. In addition, the consolidation of control service on a single channel for a relatively large area is operationally attractive because it reduces channel switching requirements in the airborne radios. The net result of such consolidation would be to combine several existing control sectors, each with its own voice channel, into a single control area with a combined voice and data channel.

Suppose now that the control area data communication system fails and an attempt is made to revert to voice service on a single channel. It is clear that the ability of this channel to provide satisfactory control service in the voice mode is less than that of the original system with its multiple sectors and channels and that a potentially hazardous control situation may develop.

In the foregoing example, the key factor in the hazard potential is the sudden drop in communication network capacity which develops if data communication fails. The hazard may of course be considerably reduced if enough standby channels are available to the network,

or if sufficient standby equipment can be provided. It appears that an appreciable reduction of the number of radio channels allocated to the control service may not be practical even if automatic data communication on a single channel is provided.

7.3 Summary

With regard to the relative merits of PSK-AM and FSK-AM, the results presented earlier in this report indicate that the theoretical edge enjoyed by PSK-AM is relatively small in those cases where the pulse rate is low with respect to the sub-carrier frequency and requires more elaborate demodulators to be exploited in full. However, it is inherently capable of higher data rates in a voice frequency channel. The rate advantage of PSK-AM may be as much as 3:1 with very little loss in efficiency. We may conclude that PSK-AM with synchronous keying is indicated when data rates above 500 bits per second are required in a nominal 3 kc/s voice band.

The design of communications terminal equipment for air traffic control is dependent on the characteristics of the communication network which it serves. Certain basic factors such as the sequencing plan, network topology, operational requirements, and reliability have far-reaching implications in terms of both system and terminal equipment design, and the utility and flexibility of terminal equipments designed without adequate attention to these factors is likely to be limited.

SECTION 8.0 BIBLIOGRAPHY

1. Hamming, R. W., "Error Detecting and Error Correcting Codes," B.S.T.J., Vol. 29, No. 2, April 1950, p. 147.

2. Silverman, R. A., and Martin Balser. "Coding for Constant - Data - Rate Systems, Part I. A New Error Correcting Code," Proc. IRE, Vol. 42, September 1954, p. 1428.

3. Montgomery, G. F., "A Comparison of Amplitude and Angle Modulation for Narrow Band Communication of Binary Coded Messages in Fluctuation Noise," Proc. IRE, Vol. 42, No. 2, February 1954, p. 447.

4. Federal Telephone and Radio Corporation, "Reference Data for Radio Engineers," 1957.

5. French, J. L., The Ministry of Civil Aviation, "Area Coverage Network-Audio Frequency Distrubution," Elec. Eng. (London), Vol. 23, April 1951, p. 146.

6. Taylor, D. P., "The Development of UHF Area Coverage Network for Civil Aviation," Elec. Eng. (London), Vol. 23, March 1951, p. 86.

7. Brinkley, J. R., "A Method of Increasing the Range of UHF Communication Systems by Multi-Carrier Amplitude Modulation," Journal IEE (Britain), Vol. 93, Part III, p. 159.

8. Norton, K. A., and P. L. Rice, "Gapless Coverage in Air-to-Ground Communication at Frequencies above 50 Mc/s," Proc. IRE, Vol. 40, No. 4, April 1952, p. 470.

9. Scholes, D. H. C., The Ministry of Civil Aviation, "UHF Area Coverage Network: Provision of Transmitting Station Equipment," Elec. Eng. (London), Vol. 23, April 1951, p. 148.

10. Law, H. B., "The Signal/Noise Performance Rating of Receivers for Long Distance Synchronous Radiotelegraph Systems Using Frequency Modulation," Proc. IEE, Vol. 104, Part B, 1959, p. 124.

11. Goldman, Stanford, "Frequency Analysis, Modulation, and Noise," McGraw-Hill, 1948.

12. Kotel'nikov, V. A., "The Theory of Optimum Noise Immunity," McGraw-Hill, 1960.

13. Bloom, F. ., "Binary Transmission by Null Zone Reception," Proc. IRE, Vol. 45, No. 7, July 1957, p. 963.

14. American Telephone and Telegraph Company, "Principles of Electricity Applied to Telephone and Telegraph Work," 1953, p. 160.

15. Daystrom Systems, "High Speed Communications Equipment," Final Report for the Federal Aviation Agency Research and Development Bureau, Study Contract FAA/BRD-118.

16. Wood, F. B., "Optimum Block Length for Data Transmission with Error Checking," Communications and Electronics, No. 40, January 1959, p. 855.

17. Marchand, N., "Air-Ground Data Transfer-Available Techniques," presented at the symposium on Air-Ground Data Transfer, RTCA

Fall Assembly Meeting, 1960.

18. Proceedings of the National Electronics Conference, Vol. XVI, 1960, pp. 1 66 and 220-250.

19. Metzner, J. J., and K. C. Morgan, "Coded Binary Decision - Feedback Communication System," IRE Transactions on Communications Systems, Vol. CS-8, No. 2, June 1960, p. 101.

20. Glenn, A. B., "Comparison of PSK vs. FSK and PSK-AM vs. FSK-AM Binary Coded Transmission Systems," IRE Transactions on Communications Systems, Vol. CS-8, No. 2, June 1960, p. 87.

21. Barker, R. H., "Group Synchronizing of Binary Digital Systems," Communication Theory, edited by Willis Jackson, Academic Press Inc., 1953.

22. Mauley, J. M., "Synchronization for the PCM Receiver," Bell Laboratories Record, Vol. 27, February 1949, p. 62.

23. Turner, F. T., "Communications Synchronizing Systems," Elec. Eng., Vol. 72, October, 1953, p. 874.

24. Edson, J. O., M. A. Flavin, and A. D. Perry, "Synchronized Clocks for Data Transmissions," Trans. AIEE, Part I, Vol. 77, January 1959, p. 832.

25. Gilbert, E. N., "Synchronization of Binary Messages," IRE Transactions on Information Theory, IT-6, No. 4, September 1960, p. 470.

26. "Telemetry System Study," Aeronutronic Publication V-743, Vol. I, pp. 3-23, ASTIA No. AD 234 958.

27. Fleck, P. L., Jr., "Transistorized Matched Filter for Pulsed Sinusoids," MIT Lincoln Lab. Report No. 156.

28. "Analysis of Advanced Data Transmission Techniques," Stromberg-Carlson Report, Contract No. FAA/BRD-80, March 1960.

29. Caldwell, S. H., "Switching Circuits and Logical Design," John Wiley and Sons, New York 1958.

30. Smith, F. Langford, "Radiotron Designer's Handbook," 1953, p. 1088 and ff.

31. McCoy, R. E., "F. M. Transient Response of Band Pass Circuits," Proc. IRE, Vol. 42, No. 3, March 1954, p. 574.